非线性最优化理论与方法

Nonlinear Optimization Theory and Methods

谢 政 李建平 陈 挚 编著

高等教育出版社



ISBN 978-7-04-028076-0

定价 34.40 元

非线性最优化理论与方法

Feixianxing Zuiyouhua Lilun yu Fangfa



高等教育出版社·北京
HIGHER EDUCATION PRESS BEIJING

内容简介

本书全面而系统地介绍了非线性最优化的理论与方法，注重基础、突出应用、强调数学建模和计算机应用能力。选材丰富，算法齐全，编排科学，论述严谨。

内容包括最优化问题的建模、无约束最优化和约束最优化问题的理论和各种算法，以及二次规划、凸规划和线性分式规划的一些特殊算法。只需具备微积分和线性代数的知识即可读懂本书。

本书既可以作为数学、信息科学、运筹学、管理科学、系统科学、控制论、计算机科学和工程技术等专业高年级本科生和研究生的教材，也可以作为相关专业的学者和技术人员的参考书。

图书在版编目(CIP)数据

非线性最优化理论与方法／谢政，李建平，陈挚编著.
—北京：高等教育出版社，2010.1
ISBN 978-7-04-028076-0

Ⅰ.非… Ⅱ.①谢…②李…③陈… Ⅲ.非线性-最优化算法-高等学校-教材 Ⅳ.O224

中国版本图书馆 CIP 数据核字（2009）第 224405 号

策划编辑 张长虹　责任编辑 张长虹　封面设计 李卫青
责任绘图 黄建英　版式设计 史新薇　责任校对 胡晓琪
责任印制 毛斯璐

出版发行	高等教育出版社	购书热线	010-58581118
社　　址	北京市西城区德外大街4号	咨询电话	400-810-0598
邮政编码	100120	网　　址	http://www.hep.edu.cn
总　　机	010-58581000		http://www.hep.com.cn
		网上订购	http://www.landraco.com
经　　销	蓝色畅想图书发行有限公司		http://www.landraco.com.cn
印　　刷	北京中科印刷有限公司	畅想教育	http://www.widedu.com
开　　本	787×960　1/16	版　　次	2010年1月第1版
印　　张	21.75	印　　次	2010年1月第1次印刷
字　　数	400 000	定　　价	34.40元

本书如有缺页、倒页、脱页等质量问题，请到所购图书销售部门联系调换。

物料号 28076-00

前　言

从古到今，"最优化"无处不在，大到人类认识世界、改造世界，小到个人理财、时间安排等生活的方方面面。最优化的思想起源于远古，最优化作为一门学科则形成于20世纪30年代。随着科学技术尤其是计算机技术的高速发展，最优化理论与方法已经在21世纪这个信息时代起着越来越重要的作用。

本书结合作者多年的教学体会与心得，本着加强最优化方法的基础理论、突出非线性最优化的应用背景、提高数学建模及计算机应用能力的原则，参照非线性最优化的最新发展，较全面、系统地介绍了非线性最优化的理论与方法。

全书共分十二章：第一章绪论，通过大量的实例的模型建立、修改，介绍最优化问题的应用背景和建模思路，引出最优化各分支，并照应各个不同专业的知识结构，将线性代数与微积分结合起来建立向量微分学概念；第二章凸性和第三章最优性条件是非线性最优化的理论基础；第四章用非线性观点讲述线性规划内容，既保证了知识的完整性，也使熟悉线性规划的读者能有新的理解，强调了线性规划与非线性规划的内在统一性；第五章迭代算法和第六章一维搜索是非线性最优化算法设计与性能分析的基础；第七章无约束最优化的解析法、第八章无约束最优化的直接法、第九章可行方向法和第十章罚函数法与广义乘子法，较全面、系统地介绍了无约束最优化和约束最优化问题的各种算法，并分析了这些算法的性能特点，给出了收敛性和收敛速度的理论证明；最后两章介绍了应用广泛、结构相对简单的二次规划、凸规划和线性分式规划的一些特殊算法。只需具备微积分和线性代数的知识即可读懂本书。

本书第二章、第三章、第四章、第七章、第八章和第十章由谢政执笔，第一章、第五章、第六章、第九章和第十二章由李建平执笔，第十一章由陈挚执笔。

本书的撰写和出版得到了许多同仁的关心和支持，得到了高等教育出版社研究生分社王瑜社长和张长虹编辑的帮助，在此表示诚挚的谢意。

作　者

2009年3月

目　录

第一章　绪论

最优化是一门应用性强、内容丰富的年轻学科，它研究最优化问题的最优解，即针对给出的实际问题，从众多的可行方案中选出最优方案.

本章包含模型与实例、数学预备知识、最优化问题的图解法等内容.

1.1　模型与实例

本节结合几个实例的数学模型的建立和修改，介绍最优化问题的应用背景和建立模型的思路，引出最优化的各个分支.

1.1.1　什么是最优化

最优化(optimization)是一门应用相当广泛的学科，它讨论决策问题的最佳选择之特性，构造寻求最佳解的计算方法，研究这些计算方法的理论性质及实际计算表现. 最优化问题广泛见于工程设计、经济规划、生产管理、交通运输、国防等重要领域. 例如，在工程设计中，怎样选择设计参数，使得设计方案既满足设计要求，又能降低成本；在资源分配中，怎样分配有限资源，使得分配方案既能满足各方面的基本要求，又能获得好的经济效益；在生产计划安排中，选择怎样的计划方案才能提高产值和利润；在原料配比问题中，确定怎样的比例才能提高质量，降低成本；在城建规划中，怎样安排布局才能有利于城市发展；在区域经济规划中，如何发挥地区优势，挖掘潜力，发展生产力；在作战指挥中，如何合理运用火力，制订作战方案，使之能有效地消灭敌人，保存自己，等等.

最优化既是一个古老的课题，又是一门年轻的学科. 早在 17 世纪，Newton 和 Leibniz 发明微积分的时代，已经提出函数的极值问题，后来又出现了 Lagrange 乘子法，Cauchy 的最速下降法. 但直到 20 世纪 30 年代，最优化的理论和方法才得以迅速发展，并不断完善，逐步成为一门系统的学科. 1939 年 Kantorovich 和 Hitchcock 等人在生产组织管理和制定交通运输方案方面首先研究和应用了线性规划. 1947 年，Dantzig 提出了求解线性规划的单纯形法，为线性规划的理论和算法奠定了基础，单纯形法被誉为“20 世纪最伟大的创造之一”. 1951 年，由 Kuhn 和 Tucker 完成了非线性规划的基础性工作. 到 20 世纪 70 年

代，最优化无论在理论和算法上，还是在应用的深度和广度上都有了进一步的发展. 随着计算机的飞速发展，最优化的应用能力越来越强，从而成为一门十分活跃的学科.

1.1.2 经典极值问题的一个实例

下面，我们用一个简单实例来回顾一下经典极值问题.

把一个半径为 1 的实心金属球熔化后，铸成一个实心圆柱体，问圆柱体取什么尺寸才能使它的表面积最小?

这是一个求圆柱体表面积受限制于(subject to，简记为 s. t.)体积为定数的约束条件的函数极值问题. 设圆柱体的底面半径为 r，高为 h，则该问题可表达为

$$\begin{cases}\min\ 2\pi rh + 2\pi r^2;\\ \text{s. t.}\ \ \pi r^2 h = \dfrac{4}{3}\pi.\end{cases}$$

也可化为无约束的函数极值问题

$$\min \frac{8\pi}{3r} + 2\pi r^2.$$

这个简单的例子，实际上代表了经典最优化中两种类型的问题.

第一，无约束极值问题

$$\min f(x_1, x_2, \cdots, x_n),$$

其中 $f(x_1, x_2, \cdots, x_n)$ 为 $\mathbb{R}^n$ 上的可微函数. 求极值点的方法是：先求出如下 n 元非线性方程组

$$\begin{cases}f'_{x_1}(x_1, x_2, \cdots, x_n) = 0,\\ f'_{x_2}(x_1, x_2, \cdots, x_n) = 0,\\ \cdots\cdots\cdots\cdots\\ f'_{x_n}(x_1, x_2, \cdots, x_n) = 0\end{cases}$$

的解，然后判定或验证这些解是不是极值点.

第二，具有等式约束的极值问题

$$\begin{cases}\min f(x_1, x_2, \cdots, x_n);\\ \text{s. t.}\ h_j(x_1, x_2, \cdots, x_n) = 0, j = 1, 2, \cdots, l,\end{cases}$$

通常用 Lagrange 乘子法(Lagrange multiplier method)来求解，即把问题转化为求 Lagrange 函数

$$L(x_1, x_2, \cdots, x_n; v_1, v_2, \cdots, v_l) = f(x_1, x_2, \cdots, x_n) - \sum_{j=1}^{l} v_j h_j(x_1, x_2, \cdots, x_n)$$

的无约束极值问题.

众所周知，上述极值问题的求解都归结为非线性方程组的求解，只有特殊情

况才能手算出来.

计算机的产生和发展,为人们求解大规模最优化应用问题提供了有用的工具,最优化理论和方法才得以迅速发展. 与经典最优化相比,近代最优化无论是模型还是求解方法都要丰富得多.

1.1.3 最优化问题的模型及分类

最优化问题所对应的模型具有如下结构:

第一是决策变量(decision variable),即所考虑的问题可归结为优选若干个被称为参数或变量的量 $x_1, x_2, \cdots, x_n$,它们都取实数值,它们的一组值构成了一个方案.

第二是约束条件(constraint condition),即对决策变量 $x_1, x_2, \cdots, x_n$ 所加的限制条件,通常用不等式和等式表示为

$$g_i(x_1, x_2, \cdots, x_n) \geqslant 0, i = 1,2,\cdots,m,$$
$$h_j(x_1, x_2, \cdots, x_n) = 0, j = 1,2,\cdots,l.$$

第三是目标函数(objective function)和目标,如使利润达到最大或使成本达到最小,通常刻画为极大化(maximize)或极小化(minimize)一个实值函数 $f(x_1, x_2, \cdots, x_n)$.

因此,最优化问题可理解为确定一组决策变量在满足约束条件下寻求目标函数的最优.

注意到极大化目标函数 $f(x_1, x_2, \cdots, x_n)$ 相当于极小化 $-f(x_1, x_2, \cdots, x_n)$. 采用向量记法,令 $\boldsymbol{x} = (x_1, x_2, \cdots, x_n)^{\mathrm{T}}$,并将约束条件写成集约束的形式,即令

$$S = \{\boldsymbol{x} \mid g_i(\boldsymbol{x}) \geqslant 0, i = 1,2,\cdots,m; h_j(\boldsymbol{x}) = 0, j = 1,2,\cdots,l\},$$

则最优化问题一般地可表述为如下形式:

$$\begin{cases} \min\ f(\boldsymbol{x}); \\ \text{s.t.}\ \ \boldsymbol{x} \in S, \end{cases} \tag{1.1.1}$$

其中称 $\boldsymbol{x} = (x_1, x_2, \cdots, x_n)^{\mathrm{T}} \in \mathbb{R}^n$ 为决策变量,$f(\boldsymbol{x})$ 为目标函数,$S \subseteq \mathbb{R}^n$ 为约束集(constraints set)或可行域(feasible region),它是所有可行解(feasible solution)即满足约束条件的点的集合.

特别地,如果可行域 $S = \mathbb{R}^n$,则最优化问题(1.1.1)称为无约束最优化问题(unconstrained optimization problem),其形式通常为

$$\min f(\boldsymbol{x}), \boldsymbol{x} \in \mathbb{R}^n. \tag{1.1.2}$$

一般地,约束最优化问题(constrained optimization problem)写作

$$\begin{cases} \min & f(\boldsymbol{x}); \\ \text{s.t.} & g_i(\boldsymbol{x}) \geqslant 0, i = 1,2,\cdots,m, \\ & h_j(\boldsymbol{x}) = 0, j = 1,2,\cdots,l, \end{cases} \tag{1.1.3}$$

其中称 $g_i(\boldsymbol{x})(i=1,2,\cdots,m),h_j(\boldsymbol{x})(j=1,2,\cdots,l)$ 为约束函数(constraint function),$g_i(\boldsymbol{x})\geqslant 0(i=1,2,\cdots,m)$ 为不等式约束,$h_j(\boldsymbol{x})=0(j=1,2,\cdots,l)$ 为等式约束. 当目标函数和约束函数均为变量 $\boldsymbol{x}$ 的线性函数时,问题(1.1.3)称为线性规划问题(linear programming problem). 当目标函数和约束函数中至少有一个是变量 $\boldsymbol{x}$ 的非线性函数时,问题(1.1.3)称为非线性规划问题(nonlinear programming problem). 此外,根据决策变量、目标函数和约束的不同,最优化还可分成整数规划(integer programming)、动态规划(dynamical programming)、网络优化(network optimization)、非光滑优化(non-smooth optimization)、随机规划(stochastic programming)、几何规划(geometric programming)、目标规划(goal programming)、模糊规划(fuzzy programming)等若干分支. 本书主要介绍求解无约束最优化问题(1.1.2)和约束最优化问题(1.1.3)的理论和方法.

1.1.4 最优化问题举例

(1) 多参数曲线拟合问题

已知热敏电阻 R 依赖于温度 t 的函数关系为

$$R(t)=x_1\exp\left(\frac{x_2}{t+x_3}\right), \tag{1.1.4}$$

其中 x_1,x_2 和 x_3 是待定的参数. 通过实验,测得一组数据 $\{(t_i,R_i)\mid i=1,2,\cdots,n\}$. 问题是如何确定参数 x_1,x_2 和 x_3?

可以设想,任意给定参数 x_1,x_2,x_3 的一组数值,就由(1.1.4)式确定了 R 关于 t 的一个函数关系式. 在几何上,它对应于一条曲线,这条曲线不一定正好通过那些测量点,一般都要产生"偏差"(如图 1.1.1 所示). 通常用最小平方和误差来度量曲线拟合的好坏,即误差 $\sum_{i=1}^{n}[R_i-R(t_i)]^2$ 越小,说明曲线拟合得越好,参数选得越合理. 于是,问题转化为求解无约束最优化问题

$$\min f(x_1,x_2,x_3)=\sum_{i=1}^{n}[R_i-R(t_i)]^2.$$

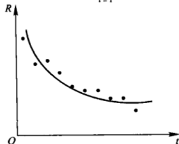


图 1.1.1 曲线拟合

这是一个典型的非线性最小二乘问题.

(2) 生产成本问题

在数量经济学中,常常用生产函数来描述经济行为的规律性,设 x_1 是作为资本的货物,x_2 为劳动力,则有著名的 Cobb - Douglas 生产函数

$$Q(x_1, x_2) = Ax_1^{\alpha}x_2^{\beta},$$

其中 Q 为产出产量,A 为生产技术水平,α 和 β 为参数. 已知工资率为 w,资本报酬率为 r,则由经济学可知生产成本为 $C = rx_1 + wx_2$. 生产成本问题是在产量不低于某水平 Q_0 的条件下,极小化生产成本. 它对应的数学模型为

$$\begin{cases} \min\ C = rx_1 + wx_2; \\ \text{s.t.}\ Ax_1^{\alpha}x_2^{\beta} \geqslant Q_0, \\ \qquad x_1 \geqslant 0, \\ \qquad x_2 \geqslant 0. \end{cases}$$

这是一个约束最优化问题,根据其特殊性,也可视为几何规划问题,其目标函数和约束函数具有多元多项式形式.

(3) 结构设计问题

结构设计师的传统工作在于力图提出这样的设计,该设计能够安全地承受所设计的载荷,最优性的概念仅仅隐含在设计的标准规范和设计师的经验之中. 但是现在有了计算机的帮助,复杂的结构设计如航空空间结构设计,对最优性的要求更明显了.

以两杆对称桁架的极小重量设计为例. 考虑图 1.1.2 所示的平面桁架,它由两根钢管构成,顶点为两杆的共同端点,两杆的另两个支点固定. 已知桁架的跨度为 $2L$,承受负荷为 $2P$,钢管的厚度 T,材料密度为 ρ,纵向弹性模数为 E,容许应力为 σ. 要确定钢管的平均直径 d 和桁架高度 h,使桁架重量最小.

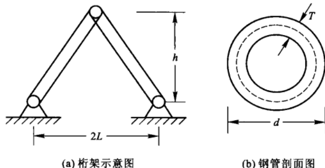


图 1.1.2 两杆对称桁架极小重量设计

这里,决策变量为 d 和 h,目标函数为

$$W = 2\rho\pi Td\sqrt{L^2 + h^2},$$

约束条件包括：

① 钢管内压应力不超过容许极限，即

$$\frac{P\sqrt{L^2 + h^2}}{\pi Tdh} \leqslant \sigma;$$

② 为保证在负荷 $2P$ 的作用下钢管不发生弯曲，要求压应力不超过临界应力 σ_l，由材料力学知

$$\sigma_l = \frac{\pi^2 E(d^2 + T^2)}{8(L^2 + h^2)} \geqslant \frac{P\sqrt{L^2 + h^2}}{\pi dhT};$$

③ 设计变量 d 和 h 有界.

于是，我们得到该问题的最优化模型为非线性规划问题

$$\begin{cases} \min\ 2\pi\rho dT\sqrt{L^2 + h^2}; \\ \text{s.t.}\ \sigma - \dfrac{P\sqrt{L^2 + h^2}}{\pi dhT} \geqslant 0, \\ \qquad \dfrac{\pi^2 E(d^2 + T^2)}{8(L^2 + h^2)} - \dfrac{P\sqrt{L^2 + h^2}}{\pi dhT} \geqslant 0, \\ \qquad d_{\max} \geqslant d \geqslant d_{\min} \\ \qquad h_{\max} \geqslant h \geqslant h_{\min}. \end{cases}$$

(4) 海洋运输问题

某航运公司承接了一项将客户停放在港口等待运输的 N 种货物运往目的地的业务. 设航运公司运输单位货物 i 的收益为 c_i(元/吨)，货船能够装载的货物的重量限制为 W(吨)，相应的容积限制为 V(立方米). 设 a_i 是单位货物 i 所占的容积(立方米/吨)，b_i 是货物 i 可提供的最大数量(吨). 问如何确定货船的装载方案，使航运公司获利最大?

设 $x_i(i=1,2,\cdots,N)$ 是货船装载货物 i 的数量，则容易建立如下变量有界的线性规划模型

$$\begin{cases} \max\ \sum\limits_{i=1}^{N} c_i x_i; \\ \text{s.t.}\ \sum\limits_{i=1}^{N} x_i \leqslant W, \\ \qquad \sum\limits_{i=1}^{N} a_i x_i \leqslant V, \\ \qquad 0 \leqslant x_i \leqslant b_i, i = 1,2,\cdots,N. \end{cases}$$

可能会发现，求解这一模型所获得的最优解并不符合实际，这就要求对模型进行修改.

下面来分析目标函数是否合理.

我们知道，每种货物的装船速度不一定相同，装船速度决定了货物在港口泊位的时间. 泊位的时间越长，泊位的费用就越大. 因此，建模时应把泊位时间因素或泊位费用考虑进去.

设货物 i 的日平均装船速度为 w_i 吨，货船的日泊位费为 q_1，货船在海上航行时的日费用为 q_2，同时假定航行距离为 d，航行速度为 v，那么货船的旅程时间为装船时间和航行时间两部分之和，表示为

$$T = \sum_{i=1}^{N} \frac{x_i}{w_i} + \frac{d}{v},$$

货船从该项运输服务中所获得的净收益为

$$P = \sum_{i=1}^{N} c_i x_i - \sum_{i=1}^{N} \frac{q_1 x_i}{w_i} - \frac{q_2 d}{v}.$$

即使把净收益最大当做评价标准有时也是不恰当的. 因为装载货物不同，货船的旅程时间 T 未必相同. 这就可能导致一种不合理的决策：在长时间的旅程中获得了一个较大的净收益. 为此考虑将单位时间内的平均利润即净收益率最大作为目标，得到修改的模型为

$$\begin{cases} \max\ z = \dfrac{P}{T}; \\ \text{s.t.}\ \ \displaystyle\sum_{i=1}^{N} x_i \leqslant W, \\ \qquad \displaystyle\sum_{i=1}^{N} a_i x_i \leqslant V, \\ \qquad 0 \leqslant x_i \leqslant b_i, i = 1,2,\cdots,N. \end{cases}$$

这是一类特殊的非线性规划问题，其中约束函数均为线性函数，只有目标函数为两个线性函数之比的非线性函数，称之为线性分式规划(linear fractional programming)问题.

(5) 资源分配问题

考虑将 m 种资源安排给 n 种活动，问应如何分配资源才能使收益最大？

设决策变量 $x_j (j=1,2,\cdots,n)$ 表示选用活动 j 的水平，已知的数据是：

$\boldsymbol{b}=(b_1,b_2,\cdots,b_m)^{\mathrm{T}}$，其中 b_i 表示第 i 种资源的拥有量；

$\boldsymbol{A}=(a_{ij})_{m\times n}$，其中 a_{ij} 表示资源 i 相对于活动 j 的单位消耗量；

$\boldsymbol{c}=(c_1,c_2,\cdots,c_n)^{\mathrm{T}}$，其中 c_j 表示活动 j 的单位利润，

则对应的最优化模型为线性规划问题

$$\begin{cases} \max\ z = \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}; \\ \text{s.t.}\ \ \boldsymbol{A}\boldsymbol{x} \leqslant \boldsymbol{b}, \\ \qquad \boldsymbol{x} \geqslant \boldsymbol{0}. \end{cases} \tag{1.1.5}$$

在某些实际应用中,上述确定型决策模型是不适宜的,因为利润系数 c_1, $c_2,\cdots,c_n$ 并不是固定的数值,而是随机变量.假定 $\boldsymbol{c}$ 是一个均值为 $\bar{\boldsymbol{c}}=(\bar{c}_1,\bar{c}_2,\cdots,\bar{c}_n)^{\mathrm{T}}$、协方差矩阵为 $\boldsymbol{V}$ 的随机向量,那么问题(1.1.5)中的目标函数 z 就是一个均值为 $\bar{\boldsymbol{c}}^{\mathrm{T}}\boldsymbol{x}$、方差为 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{V}\boldsymbol{x}$ 的随机变量.

① 如果希望 z 的期望值最大,则可考虑线性规划模型

$$\begin{cases}\max\ \bar{\boldsymbol{c}}^{\mathrm{T}}\boldsymbol{x};\\ \text{s.t.}\ \ \boldsymbol{A}\boldsymbol{x}\leqslant\boldsymbol{b},\\ \qquad\ \boldsymbol{x}\geqslant\boldsymbol{0}.\end{cases}$$

② 如果希望 z 的方差最小,则有如下模型

$$\begin{cases}\min\ \boldsymbol{x}^{\mathrm{T}}\boldsymbol{V}\boldsymbol{x};\\ \text{s.t.}\ \ \boldsymbol{A}\boldsymbol{x}\leqslant\boldsymbol{b},\\ \qquad\ \boldsymbol{x}\geqslant\boldsymbol{0}.\end{cases}\tag{1.1.6}$$

问题(1.1.6)中,目标函数是二次函数(quadratic function),约束均为线性约束,它是一类特殊的非线性规划问题,称之为二次规划(quadratic programming)问题.

③ 如果希望期望最大且方差最小,则得到一个多目标规划(multiple objective programming)模型

$$\begin{cases}\max\ \bar{\boldsymbol{c}}^{\mathrm{T}}\boldsymbol{x};\\ \min\ \boldsymbol{x}^{\mathrm{T}}\boldsymbol{V}\boldsymbol{x};\\ \text{s.t.}\ \ \boldsymbol{A}\boldsymbol{x}\leqslant\boldsymbol{b},\\ \qquad\ \boldsymbol{x}\geqslant\boldsymbol{0}.\end{cases}$$

④ 对期望和方差两者加以综合考虑.

假定人们的兴趣在于保证期望值至少达到某一值 $\bar{z}$,这个值 $\bar{z}$ 常常被认为是愿望水平或满意水平.

一种自然的考虑方法是要求 $\bar{\boldsymbol{c}}^{\mathrm{T}}\boldsymbol{x}\geqslant\bar{z}$,那么,我们得到如下模型

$$\begin{cases}\min\ \boldsymbol{x}^{\mathrm{T}}\boldsymbol{V}\boldsymbol{x};\\ \text{s.t.}\ \ \boldsymbol{A}\boldsymbol{x}\leqslant\boldsymbol{b},\\ \qquad\ \bar{\boldsymbol{c}}^{\mathrm{T}}\boldsymbol{x}\geqslant\bar{z},\\ \qquad\ \boldsymbol{x}\geqslant\boldsymbol{0}.\end{cases}$$

这仍然是一个二次规划问题.

另一种考虑方法是令 $\alpha=\mathrm{Prob}(\boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}\geqslant\bar{z})$,$\alpha$ 是获得满意水平 $\bar{z}$ 的概率,当然人们喜欢极大化 α.为转化为确定型问题,设 $\boldsymbol{c}=\boldsymbol{d}+y\boldsymbol{p}$,其中 $\boldsymbol{d},\boldsymbol{p}$ 是 n 维固定向量,y 为一维随机变量,则

$$\alpha=\mathrm{Prob}(\boldsymbol{d}^{\mathrm{T}}\boldsymbol{x}+y\boldsymbol{p}^{\mathrm{T}}\boldsymbol{x}\geqslant\bar{z})=\mathrm{Prob}\left(y\geqslant\frac{\bar{z}-\boldsymbol{d}^{\mathrm{T}}\boldsymbol{x}}{\boldsymbol{p}^{\mathrm{T}}\boldsymbol{x}}\right),$$

相应的模型为如下线性分式规划问题

$$\begin{cases} \min \ \dfrac{\bar{z} - \boldsymbol{d}^{\mathrm{T}}\boldsymbol{x}}{\boldsymbol{p}^{\mathrm{T}}\boldsymbol{x}}; \\ \text{s.t.} \ \ \boldsymbol{A}\boldsymbol{x} \leqslant \boldsymbol{b}, \\ \qquad \boldsymbol{x} \geqslant \boldsymbol{0}. \end{cases}$$

(6) 设施问题

在实际中经常遇到活动中心的最优位置问题,如工厂或仓库的位置,工厂中机械加工中心或部门的位置,城区中应急设施(如消防、公安、救护中心等)的位置.下面考虑一种简单情况.

设有 n 个市场,第 j 个市场的位置为 (a_j, b_j),对某种货物的需求量为 $q_j, j = 1,2,\cdots,n$.现计划建立 m 个仓库,第 i 个仓库的容量为 $c_i, i = 1,2,\cdots,m$.试确定仓库的位置,使各仓库到各市场的总加权距离最小.

设 (x_i, y_i) 为仓库 i 的位置,d_{ij} 为仓库 i 到市场 j 的距离,w_{ij} 为仓库 i 到市场 j 的货物单位数量,$i = 1,2,\cdots,m; j = 1,2,\cdots,n$,则该问题的模型是

$$\begin{cases} \min \ \sum\limits_{i=1}^{m}\sum\limits_{j=1}^{n} w_{ij}d_{ij}; \\ \text{s.t.} \ \ \sum\limits_{j=1}^{n} w_{ij} \leqslant c_i, i = 1,2,\cdots,m, \\ \qquad \sum\limits_{i=1}^{m} w_{ij} = q_j, j = 1,2,\cdots,n, \\ \qquad w_{ij} \geqslant 0, i = 1,2,\cdots,m; j = 1,2,\cdots,n. \end{cases} \tag{1.1.7}$$

关于距离可采用不同的度量,如

$$d_{ij} = |x_i - a_j| + |y_i - b_j| \text{ 或 } d_{ij} = \sqrt{(x_i - a_j)^2 + (y_i - b_j)^2}.$$

这样,问题(1.1.7)就是一个以 $x_1, x_2, \cdots, x_m, y_1, y_2, \cdots, y_m, w_{11}, \cdots, w_{mn}$ 为变量的非线性规划问题.如果仓库的位置是固定的,则 d_{ij} 为已知,那么问题(1.1.7)就是众所周知的运输问题的线性规划模型.

(7) 政治区划

假设 m 个基本人口单元(如县、社团等)按某种政治需要可以得到 n 个政区组合.设 c_j 为选取政区 j 的费用或不可接受性的度量.问应怎样把这 m 个人口单元划分成 K 个政区,使总的费用(或不可接受性)最小?

设

$$x_j = \begin{cases} 1, \text{选取政区 } j, \\ 0, \text{否则}, \end{cases} \qquad j = 1,2,\cdots,n,$$

$$a_{ij} = \begin{cases} 1, \text{人口单元 } i \text{ 划到政区 } j \text{ 中}, \\ 0, \text{否则}, \end{cases} \qquad i = 1,2,\cdots,m, j = 1,2,\cdots,n,$$

则问题可表述为如下 0－1 整数规划问题

$$\begin{cases} \min & \sum_{j=1}^{n} c_j x_j; \\ \text{s.t.} & \sum_{j=1}^{n} a_{ij} x_j = 1, i = 1,2,\cdots,m, \\ & \sum_{j=1}^{n} x_j = K, \\ & x_j = 0 \text{ 或 } 1, j = 1,2,\cdots,n. \end{cases}$$

(8) 最优控制

如果最优化问题的解不随时间而变,则称为静态最优化(参数最优化)问题.前面列举的例子都属于静态最优化问题.

如果最优化问题的解随时间变化,即变量是时间 t 的函数,则称为动态最优化,也即最优控制(optimal control)问题.

一个离散控制问题可以表述为一个非线性规划问题,而一个连续最优控制问题又可以用一个非线性规划问题去逼近,因此,非线性规划问题的方法可以用来求解某些最优控制问题.

1) 离散最优控制

考虑一个持续 K 时期的固定时间的离散最优控制问题.在时期 k 开始,该系统用状态向量 $\boldsymbol{y}_{k-1}$ 表示,在时期 k 末,控制向量 $\boldsymbol{u}_k$ 按照下面的关系把系统的状态从 $\boldsymbol{y}_{k-1}$ 改变到 $\boldsymbol{y}_k$:

$$\boldsymbol{y}_k = \boldsymbol{y}_{k-1} + \boldsymbol{\Phi}_k(\boldsymbol{y}_{k-1}, \boldsymbol{u}_k), \tag{1.1.8}$$

称(1.1.8)式为状态方程.给出初始状态 $\boldsymbol{y}_0$,应用控制序列 $\boldsymbol{u}_1, \boldsymbol{u}_2, \cdots, \boldsymbol{u}_k$ 能够得到一状态向量序列 $\boldsymbol{y}_1, \boldsymbol{y}_2, \cdots, \boldsymbol{y}_k$,这些状态向量称之为轨道.这个过程如图 1.1.3 所示.

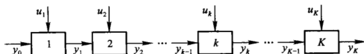


图 1.1.3 离散最优控制

控制序列 $\boldsymbol{u}_1, \boldsymbol{u}_2, \cdots, \boldsymbol{u}_k$ 和状态向量序列 $\boldsymbol{y}_0, \boldsymbol{y}_1, \cdots, \boldsymbol{y}_k$ 称为可行的是指它们满足以下约束:

$$\boldsymbol{y}_k \in Y_k, k = 1,2,\cdots,K,$$
$$\boldsymbol{u}_k \in U_k, k = 1,2,\cdots,K,$$
$$\boldsymbol{\Psi}(\boldsymbol{y}_0, \boldsymbol{y}_1, \cdots, \boldsymbol{y}_k; \boldsymbol{u}_1, \boldsymbol{u}_2, \cdots, \boldsymbol{u}_k) \in D,$$

其中 $Y_1, Y_2, \cdots, Y_k, U_1, U_2, \cdots, U_k$ 及 D 都是给定的集合,$\boldsymbol{\Psi}$ 是已知函数,通常称为

轨道约束函数.离散最优控制问题是在所有可行控制和可行轨道中,通过优化某一目标函数来寻求一最优控制和一个相应的最优化轨道,其对应的模型是

$$\begin{cases}\min\ \alpha(\boldsymbol{y}_0,\boldsymbol{y}_1,\cdots,\boldsymbol{y}_k;\boldsymbol{u}_1,\boldsymbol{u}_2,\cdots,\boldsymbol{u}_k);\\ \text{s.t.}\ \ \boldsymbol{y}_k=\boldsymbol{y}_{k-1}+\boldsymbol{\Phi}_k(\boldsymbol{y}_{k-1},\boldsymbol{u}_k),k=1,2,\cdots,K,\\ \qquad \boldsymbol{y}_k\in Y_k,k=1,2,\cdots,K,\\ \qquad \boldsymbol{u}_k\in U_k,k=1,2,\cdots,K,\\ \qquad \boldsymbol{\Psi}(\boldsymbol{y}_0,\boldsymbol{y}_1,\cdots,\boldsymbol{y}_k;\boldsymbol{u}_1,\boldsymbol{u}_2,\cdots,\boldsymbol{u}_k)\in D.\end{cases}$$

下面我们用生产与存贮的实例来具体说明离散最优控制问题.

假定某公司生产某种产品以满足某一已知的需求,并假定生产进程超过 K 期就应停止.在每个时期的开始用存贮来满足需求,然后由生产来满足,而任何时期最大的生产量受到现有生产能力的限制,因而生产量不能超过 b 单位.假设需要时,可以雇佣足够的临时工,多余时,可以解雇这些临时工.然而,为了防止劳动力大的波动,假定任何两个相邻时期劳动力差的平方与某一费用成正比.同时,把存贮从一个时期保留到另一个时期也引起了一个与之成正比的费用.问题是如何确定劳动力和存贮的水平使得既满足需求又使费用最小.

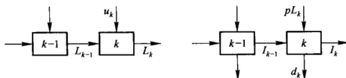


图 1.1.4 生产与存贮

如图 1.1.4 所示,在该问题中,状态变量为时期 k 末的存贮水平 I_k 和劳动力 L_k,控制变量是时期 k 中所获得的劳动力 u_k($u_k<0$ 意指解雇劳动力 $-u_k$),该问题的模型是

$$\begin{cases}\min\ \sum\limits_{k=1}^{K}(c_1u_k^{\ 2}+c_2I_k);\\ \text{s.t.}\ \ L_k=L_{k-1}+u_k,k=1,2,\cdots,K,\\ \qquad I_k=I_{k-1}+pL_k-d_k,k=1,2,\cdots,K,\\ \qquad 0\leqslant L_k\leqslant\dfrac{b}{p},k=1,2,\cdots,K,\\ \qquad I_k\geqslant 0,k=1,2,\cdots,K,\end{cases}$$

其中初始存贮量 I_0 和初始劳动力数 L_0 为已知,d_k 是时期 k 中的已知需求,p 是每个工人在任何时期中生产的产品单位数.

2）连续最优控制

在离散控制问题的情形下，控制运用在离散点上. 现在我们考虑控制函数 $\boldsymbol{u}$ 运用在一固定时间 $[0,T]$ 内的连续控制问题. 其对应的模型是

$$\begin{cases} \min \displaystyle\int_0^T \alpha[\boldsymbol{y}(t),\boldsymbol{u}(t)]\mathrm{d}t; \\ \text{s.t. } \dfrac{\mathrm{d}\boldsymbol{y}}{\mathrm{d}t} = \boldsymbol{\Phi}[\boldsymbol{y}(t),\boldsymbol{u}(t)], t \in [0,T], \\ \quad \boldsymbol{y} \in Y, \\ \quad \boldsymbol{u} \in U, \\ \quad \boldsymbol{\Psi}(\boldsymbol{y},\boldsymbol{u}) \in D, \end{cases} \tag{1.1.9}$$

其中初始状态向量 $\boldsymbol{y}(0) = \boldsymbol{y}_0$ 是已知的.

假定把时段 $[0,T]$ 分成 K 个时期，使得 $K\Delta = T$，记 $\boldsymbol{y}_k = \boldsymbol{y}(k\Delta)$，$\boldsymbol{u}_k = \boldsymbol{u}(k\Delta)$，则连续最优控制问题(1.1.9)可用下面的离散最优控制问题来逼近：

$$\begin{cases} \min \displaystyle\sum_{k=1}^{K} \alpha(\boldsymbol{y}_k,\boldsymbol{u}_k); \\ \text{s.t. } \boldsymbol{y}_k = \boldsymbol{y}_{k-1} + \boldsymbol{\Phi}_k(\boldsymbol{y}_{k-1},\boldsymbol{u}_k), k = 1,2,\cdots,K, \\ \quad \boldsymbol{y}_k \in Y_k, k = 1,2,\cdots,K, \\ \quad \boldsymbol{u}_k \in U_k, k = 1,2,\cdots,K, \\ \quad \boldsymbol{\Psi}(\boldsymbol{y}_0,\boldsymbol{y}_1,\cdots,\boldsymbol{y}_k,\boldsymbol{u}_1\cdots,\boldsymbol{u}_k) \in D. \end{cases}$$

考虑发射火箭的例子.

假设质量为 m 的火箭在时刻 t 产生的垂直向上的推力为 $u(t)$，离地面高度为 $y(t)$，并且在任何时刻火箭所产生的最大推力不超过 b. 为了耗费最小能量使火箭在时刻 T 到达高度 H，可建立如下控制模型

$$\begin{cases} \min \displaystyle\int_0^T |u(t)|\mathrm{d}t; \\ \text{s.t. } my''(t) + mg = u(t), t \in [0,T], \\ \quad |u(t)| \leqslant b, t \in [0,T], \\ \quad y(0) = 0, y(T) = H, \\ \quad y'(0) = 0, \end{cases} \tag{1.1.10}$$

其中 g 为重力加速度. 令 $y_1 = y(t)$，$y_2 = y'(t)$，可将问题(1.1.10)化为如下带一阶微分方程的等价问题

$$
\begin{cases}
\min \int_0^{T} |u(t)|\,\mathrm{d}t; \\
\text{s.t.}\ \ y'_1(t) = y_2(t), t \in [0,T], \\
\qquad y'_2(t) = \dfrac{u(t)}{m} - g, t \in [0,T], \\
\qquad |u(t)| \leqslant b, t \in [0,T], \\
\qquad y_1(0) = 0, y_1(T) = H, y_2(0) = 0.
\end{cases}
\tag{1.1.11}
$$

假定把区间$[0,T]$分成K个时期，每个时期长为1，在时期k末，推力、高度和速度分别用u_k，$y_{1,k}$和$y_{2,k}$表示，则问题(1.1.11)可用如下非线性规划模型来近似：

$$
\begin{cases}
\min \sum_{k=1}^{k} |u_k|; \\
\text{s.t.}\ \ y_{1,k} - y_{1,k-1} = y_{2,k-1}, k = 1,2,\cdots,K, \\
\qquad y_{2,k} - y_{2,k-1} = \dfrac{u_k}{m} - g, k = 1,2,\cdots,K, \\
\qquad |u_k| \leqslant b, k = 1,2,\cdots,K, \\
\qquad y_{1,0} = y_{2,0} = 0, y_{1,K} = H.
\end{cases}
$$

以上我们结合最优化问题的建模，进一步了解到最优化方法的广泛应用，同时，初步认识了最优化的一些重要模型. 它们是我们研究最优化理论与方法的基础.

1.2 数学预备知识

这一节，我们将对线性代数与微积分的基本知识作一简单回顾，以巩固向量函数微分学的有关知识.

1.2.1 向量的范数和矩阵的条件数

在本书中，约定向量取列向量形式. 这样，n 维 Euclid 空间$\mathbb{R}^n$ 中的一个向量 $\boldsymbol{x}$ 表示为

$$
\boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix},
$$

也可表示为行向量的转置，即 $\boldsymbol{x} = (x_1, x_2, \cdots, x_n)^{\mathrm{T}}$，其中 $x_1, x_2, \cdots, x_n$ 为向量 $\boldsymbol{x}$ 的分量. 分量都为 0 的向量称为零向量，记为 $\mathbf{0}$.

设 $\boldsymbol{x} = (x_1, x_2, \cdots, x_n)^{\mathrm{T}}$，$\boldsymbol{y} = (y_1, y_2, \cdots, y_n)^{\mathrm{T}}$，如果对一切 $1 \leqslant i \leqslant n$，都有 $x_i =$

y_i,则称向量 $\boldsymbol{x}$ 等于向量 $\boldsymbol{y}$,记作 $\boldsymbol{x}=\boldsymbol{y}$;如果对一切 $1\leqslant i\leqslant n$,都有 $x_i\leqslant y_i$(或 $x_i\geqslant y_i$),则称向量 $\boldsymbol{x}$ 小于等于(或大于等于)$\boldsymbol{y}$,记作 $\boldsymbol{x}\leqslant\boldsymbol{y}$(或 $\boldsymbol{x}\geqslant\boldsymbol{y}$);如果对一切 $1\leqslant i\leqslant n$,都有 $x_i<y_i$(或 $x_i>y_i$),则称向量 $\boldsymbol{x}$ 小于(或大于)$\boldsymbol{y}$,记作 $\boldsymbol{x}<\boldsymbol{y}$($\boldsymbol{x}>\boldsymbol{y}$).

在本书中,我们约定矩阵为实矩阵.把 $m\times n$ 实矩阵的全体记作$\mathbb{R}^{m\times n}$.

设 $\boldsymbol{A}$ 是 n 阶对称矩阵,$\boldsymbol{x}$ 为任意非零 n 维向量.

(1) 若 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}>0$ 则称 $\boldsymbol{A}$ 为正定矩阵;

(2) 若 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}\geqslant 0$,则称 $\boldsymbol{A}$ 为半正定矩阵;

(3) 若 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}<0$,则称 $\boldsymbol{A}$ 为负定矩阵;

(4) 若 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}\leqslant 0$,则称 $\boldsymbol{A}$ 为半负定矩阵.

除此之外,称 $\boldsymbol{A}$ 为不定矩阵.

设 $\boldsymbol{A}$ 为 n 阶对称矩阵,利用 $\boldsymbol{A}$ 的各特征值或顺序主子式的符号可以判定 $\boldsymbol{A}$ 的正定性(见表 1.2.1).

表 1.2.1　对称矩阵的正定性判定

$\boldsymbol{A}$ 为正定矩阵	$\boldsymbol{A}$ 的特征值都大于 0	$\boldsymbol{A}$ 的各阶顺序主子式都大于 0
$\boldsymbol{A}$ 为半正定矩阵	$\boldsymbol{A}$ 的特征值都大于等于 0	$\boldsymbol{A}$ 的各阶顺序主子式都大于等于 0
$\boldsymbol{A}$ 为负定矩阵	$\boldsymbol{A}$ 的特征值都小于 0	$\boldsymbol{A}$ 的奇阶顺序主子式都小于 0,且 $\boldsymbol{A}$ 的偶阶顺序主子式都大于 0
$\boldsymbol{A}$ 为半负定矩阵	$\boldsymbol{A}$ 的特征值都小于等于 0	$\boldsymbol{A}$ 的奇阶顺序主子式都小于等于 0,且 $\boldsymbol{A}$ 的偶阶顺序主子式都大于等于 0
$\boldsymbol{A}$ 为不定矩阵	$\boldsymbol{A}$ 既有正的特征值,又有负的特征值	

设 $\boldsymbol{x},\boldsymbol{y}\in\mathbb{R}^n$,称

$$\boldsymbol{x}^{\mathrm{T}}\boldsymbol{y} = \sum_{i=1}^{n} x_i y_i = \boldsymbol{y}^{\mathrm{T}}\boldsymbol{x}$$

为向量 $\boldsymbol{x}$ 与 $\boldsymbol{y}$ 的内积.称

$$\|\boldsymbol{x}\| = \left(\sum_{i=1}^{n} x_i^2\right)^{1/2} \tag{1.2.1}$$

为向量 $\boldsymbol{x}$ 的 Euclid 范数(或向量 $\boldsymbol{x}$ 的长度).范数为 1 的向量称为单位向量.因为 n 维向量与$\mathbb{R}^n$ 中的点一一对应,因此称

$$\|\boldsymbol{x}-\boldsymbol{y}\| = \left(\sum_{i=1}^{n}(x_i-y_i)^2\right)^{1/2}$$

为两点 $\boldsymbol{x},\boldsymbol{y}$ 间的距离.

关于 Euclid 范数与内积，我们提及两个重要不等式：

(1) 三角不等式

$$\|x+y\| \leqslant \|x\| + \|y\|;$$

(2) Cauchy 不等式

$$|x^{\mathrm{T}}y| \leqslant \|x\| \cdot \|y\|.$$

以上两式等号成立当且仅当 $x=\lambda y$(λ 为实数).

更一般地，对$\mathbb{R}^n$ 中元素定义的向量范数 $\|\cdot\|$ 是$\mathbb{R}^n$ 上的一个非负实值函数，它满足以下三个条件：

(1) $\forall x \in \mathbb{R}^n$，有 $\|x\| \geqslant 0$，同时当且仅当 $x=\mathbf{0}$ 时 $\|x\|=0$；

(2) $\forall x \in \mathbb{R}^n, \forall \alpha \in \mathbb{R}$，有 $\|\alpha x\| = |\alpha| \, \|x\|$；

(3) $\forall x, y \in \mathbb{R}^n$，有 $\|x+y\| \leqslant \|x\| + \|y\|$.

例如，p 范数

$$\|x\|_p = \left(\sum_{i=1}^{n} |x_i|^p\right)^{1/p} (p \geqslant 1)$$

和最大范数

$$\|x\|_\infty = \max\{|x_i| \mid 1 \leqslant i \leqslant n\}$$

都可以作为向量范数.

对$\mathbb{R}^n$ 中向量可引进各种各样的范数，但它们具有下述性质：

设 $\|\cdot\|_A$ 和 $\|\cdot\|_B$ 是定义于$\mathbb{R}^n$ 中的两种向量范数，则总存在正数 c_1 和 c_2，使对一切 $x \in \mathbb{R}^n$，有

$$c_1 \|x\|_A \leqslant \|x\|_B \leqslant c_2 \|x\|_A. \tag{1.2.2}$$

这一性质称为向量范数的等价性.

在最优化方法中，常常需要考虑某个点列 $\{x_k\}$ 的收敛快慢. 设 $x_k \to \bar{x}$，需考虑 $\|x_k - \bar{x}\|$ 趋于 0 的速率，利用向量范数的等价性(1.2.2)式，只需对某一种向量范数进行考虑即可. 本书中除特别声明外，所出现的范数都是 Euclid 范数.

下面引进矩阵的条件数的概念.

称 n 阶方阵 A 的各特征值的绝对值 $|\lambda_1|, |\lambda_2|, \cdots, |\lambda_n|$ 中的最大者为 A 的谱半径.

设 $A \in \mathbb{R}^{m \times n}$，称方阵 $A^{\mathrm{T}}A$ 的特征值的非负平方根为 A 的奇异值. 矩阵 A 的(谱)条件数定义为 A 的最大奇异值与最小非零奇异值之商，记为 $\kappa(A)$，即

$$\kappa(A) = \frac{\sigma_1(A)}{\sigma_t(A)},$$

其中

$$\sigma_1(A) \geqslant \sigma_2(A) \geqslant \cdots \geqslant \sigma_t(A) \geqslant \cdots \geqslant \sigma_n(A)$$

为 A 的所有奇异值，且 t 为 A 的秩(即 $t = \operatorname{rank} A$).

如果 $\boldsymbol{A}$ 为 n 阶正定矩阵，则记

$$\lambda_1(\boldsymbol{A}) \geqslant \lambda_2(\boldsymbol{A}) \geqslant \cdots \geqslant \lambda_n(\boldsymbol{A}) > 0$$

和

$$\sigma_1(\boldsymbol{A}) \geqslant \sigma_2(\boldsymbol{A}) \geqslant \cdots \geqslant \sigma_n(\boldsymbol{A})$$

分别为 $\boldsymbol{A}$ 的特征值和奇异值，从而 $\sigma_i(\boldsymbol{A}) = \lambda_i(\boldsymbol{A}), i = 1, 2, \cdots, n$，于是

$$\kappa(\boldsymbol{A}) = \frac{\lambda_1(\boldsymbol{A})}{\lambda_n(\boldsymbol{A})},$$

即正定矩阵的条件数等于它的最大特征值与最小特征值之商.

如果 $\boldsymbol{A}$ 为 n 阶满秩矩阵，则 $\boldsymbol{A}$ 的所有奇异值 $\sigma_1(\boldsymbol{A}) \geqslant \sigma_2(\boldsymbol{A}) \geqslant \cdots \geqslant \sigma_n(\boldsymbol{A}) > 0$，从而

$$\kappa(\boldsymbol{A}) = \frac{\sigma_1(\boldsymbol{A})}{\sigma_n(\boldsymbol{A})},$$

即满秩矩阵的条件数等于它的最大奇异值与最小奇异值之商.

一个 n 阶满秩矩阵 $\boldsymbol{A}$ 称为病态的，如果 $\boldsymbol{A}$ 的 n 个列向量之间存在着近似线性关系. 此时 $\boldsymbol{A}$ 的最小奇异值就会很小，相应地条件数就会很大. 因此条件数可以用来度量矩阵的病态程度. 例如，对于二阶方阵

$$\boldsymbol{A} = \begin{bmatrix} 1 & \varepsilon \\ \varepsilon & 1 \end{bmatrix}, 0 < \varepsilon < 1,$$

容易验证 $\sigma_1(\boldsymbol{A}) = 1 + \varepsilon, \sigma_2(\boldsymbol{A}) = 1 - \varepsilon$. 当 $\varepsilon \to 1$ 时，

$$\kappa(\boldsymbol{A}) = \frac{1 + \varepsilon}{1 - \varepsilon} \to \infty,$$

相应地，$\boldsymbol{A}$ 趋于奇异矩阵，也就是说，$\boldsymbol{A}$ 的病态程度愈来愈严重.

1.2.2 多元函数的梯度、Hesse 矩阵及 Taylor 公式

一个 n 元函数 $f(x_1, x_2, \cdots, x_n)$ 可视为向量变量 $\boldsymbol{x} = (x_1, x_2, \cdots, x_n)^{\mathrm{T}}$ 的实值函数，记作 $f(\boldsymbol{x})$.

定义 1.2.1 设 $f: \mathbb{R}^n \to \mathbb{R}, \bar{\boldsymbol{x}} \in \mathbb{R}^n$. 如果存在 n 维向量 $\boldsymbol{p}$，对任意 n 维非零向量 $\Delta \boldsymbol{x}$，使得

$$\lim_{\|\Delta \boldsymbol{x}\| \to 0} \frac{f(\bar{\boldsymbol{x}} + \Delta \boldsymbol{x}) - f(\bar{\boldsymbol{x}}) - \boldsymbol{p}^{\mathrm{T}} \Delta \boldsymbol{x}}{\|\Delta \boldsymbol{x}\|} = 0, \tag{1.2.3}$$

则称 $f(\boldsymbol{x})$ 在点 $\bar{\boldsymbol{x}}$ 处可微，并称

$$\mathrm{d} f(\bar{\boldsymbol{x}}) = \boldsymbol{p}^{\mathrm{T}} \Delta \boldsymbol{x} \tag{1.2.4}$$

为 $f(\boldsymbol{x})$ 在点 $\bar{\boldsymbol{x}}$ 处的微分.

式(1.2.3)可以写成下述等价形式

$$f(\bar{\boldsymbol{x}} + \Delta \boldsymbol{x}) - f(\bar{\boldsymbol{x}}) = \boldsymbol{p}^{\mathrm{T}} \Delta \boldsymbol{x} + o(\|\Delta \boldsymbol{x}\|). \tag{1.2.5}$$

定义 1.2.2 设$f:\mathbb{R}^n\to\mathbb{R},\bar{\boldsymbol{x}}\in\mathbb{R}^n$. 如果$f(\boldsymbol{x})$在点$\bar{\boldsymbol{x}}$处对于自变量$\boldsymbol{x}=(x_1,x_2,\cdots,x_n)^{\mathrm{T}}$的各分量的偏导数

$$\frac{\partial f(\bar{\boldsymbol{x}})}{\partial x_i},i=1,2,\cdots,n$$

都存在,则称函数f在点$\bar{\boldsymbol{x}}$处一阶可导,并且称向量

$$\nabla f(\bar{\boldsymbol{x}})=\left(\frac{\partial f(\bar{\boldsymbol{x}})}{\partial x_1},\frac{\partial f(\bar{\boldsymbol{x}})}{\partial x_2},\cdots,\frac{\partial f(\bar{\boldsymbol{x}})}{\partial x_n}\right)^{\mathrm{T}}$$

为$f(\boldsymbol{x})$在点$\bar{\boldsymbol{x}}$处的一阶导数或梯度.

定理 1.2.1 设$f:\mathbb{R}^n\to\mathbb{R},\bar{\boldsymbol{x}}\in\mathbb{R}^n$. 如果$f(\boldsymbol{x})$在点$\bar{\boldsymbol{x}}$处可微,则$f(\boldsymbol{x})$在点$\bar{\boldsymbol{x}}$处的梯度$\nabla f(\bar{\boldsymbol{x}})$存在,并且有

$$\mathrm{d}f(\bar{\boldsymbol{x}})=\nabla f(\bar{\boldsymbol{x}})^{\mathrm{T}}\Delta\boldsymbol{x}.$$

证明 在(1.2.5)式中依次取

$$\Delta\boldsymbol{x}=\Delta x_i\boldsymbol{e}_i,i=1,2,\cdots,n,$$

其中$\Delta\boldsymbol{x}=(\Delta x_1,\Delta x_2,\cdots,\Delta x_n)^{\mathrm{T}},\boldsymbol{e}_i$是第$i$个单位坐标向量(它的第$i$个分量为1,其余分量均为0),即

$$\boldsymbol{e}_i=(0,\cdots,0,1,0,\cdots,0)^{\mathrm{T}}.$$

记$\boldsymbol{p}=(p_1,p_2,\cdots,p_n)^{\mathrm{T}}$,则

$$f(\bar{\boldsymbol{x}}+\Delta x_i\boldsymbol{e}_i)-f(\bar{\boldsymbol{x}})=p_i\Delta x_i+o(|\Delta x_i|),i=1,2,\cdots,n,$$

从而

$$\lim_{\Delta x_i\to 0}\frac{f(\bar{\boldsymbol{x}}+\Delta x_i\boldsymbol{e}_i)-f(\bar{\boldsymbol{x}})}{\Delta x_i}=p_i,i=1,2,\cdots,n.$$

故$\nabla f(\bar{\boldsymbol{x}})$存在且$\boldsymbol{p}=\nabla f(\bar{\boldsymbol{x}})$,从而由(1.2.4)式知结论成立. □

根据上述定理,(1.2.5)式可写成

$$f(\bar{\boldsymbol{x}}+\Delta\boldsymbol{x})-f(\bar{\boldsymbol{x}})=\nabla f(\bar{\boldsymbol{x}})^{\mathrm{T}}\Delta\boldsymbol{x}+o(\|\Delta\boldsymbol{x}\|),\qquad(1.2.6)$$

或

$$f(\bar{\boldsymbol{x}}+\Delta\boldsymbol{x})=f(\bar{\boldsymbol{x}})+\nabla f(\bar{\boldsymbol{x}})^{\mathrm{T}}\Delta\boldsymbol{x}+o(\|\Delta\boldsymbol{x}\|).\qquad(1.2.7)$$

定义 1.2.3 设$f:\mathbb{R}^n\to\mathbb{R},\bar{\boldsymbol{x}}\in\mathbb{R}^n,\boldsymbol{d}$是给定的$n$维非零向量,$\boldsymbol{e}=\dfrac{\boldsymbol{d}}{\|\boldsymbol{d}\|}$. 如果极限

$$\lim_{\lambda\to 0}\frac{f(\bar{\boldsymbol{x}}+\lambda\boldsymbol{e})-f(\bar{\boldsymbol{x}})}{\lambda}(\lambda\in\mathbb{R})\qquad(1.2.8)$$

存在,则称此极限为$f(\boldsymbol{x})$在点$\bar{\boldsymbol{x}}$沿方向$\boldsymbol{d}$的方向导数,记作$\dfrac{\partial f(\bar{\boldsymbol{x}})}{\partial\boldsymbol{d}}$.

定理 1.2.2 设$f:\mathbb{R}^n\to\mathbb{R},\bar{\boldsymbol{x}}\in\mathbb{R}^n$. 如果$f(\boldsymbol{x})$在点$\bar{\boldsymbol{x}}$可微,则$f(\boldsymbol{x})$在$\bar{\boldsymbol{x}}$处沿任何非零向量$\boldsymbol{d}$的方向导数存在,且

$$\frac{\partial f(\bar{\boldsymbol{x}})}{\partial\boldsymbol{d}}=\nabla f(\bar{\boldsymbol{x}})^{\mathrm{T}}\boldsymbol{e},$$

其中 $e=\dfrac{d}{\|d\|}$.

证明 由(1.2.8)式和(1.2.6)式知

$$\begin{aligned}\frac{\partial f(\bar{x})}{\partial d}&=\lim_{\lambda\to 0}\frac{f(\bar{x}+\lambda e)-f(\bar{x})}{\lambda}\\&=\lim_{\lambda\to 0}\frac{\lambda\nabla f(\bar{x})^{\mathrm{T}}e+o(\lambda)}{\lambda}=\nabla f(\bar{x})^{\mathrm{T}}e.\end{aligned}$$ □

定义 1.2.4 设 $f(x)$ 是 $\mathbb{R}^n$ 上的连续函数，$\bar{x}\in\mathbb{R}^n$，d 是 n 维非零向量. 如果存在 $\delta>0$，使得

$$f(\bar{x}+\lambda d)<f(\bar{x}),\ \forall\lambda\in(0,\delta),$$

则称 d 为 f 在点 $\bar{x}$ 处的下降方向；若

$$f(\bar{x}+\lambda d)>f(\bar{x}),\ \forall\lambda\in(0,\delta),$$

则称 d 为 f 在点 $\bar{x}$ 处的上升方向.

根据(1.2.8)式，应用极限保号性定理易知，当 $\dfrac{\partial f(\bar{x})}{\partial d}<0$ 时，$f(x)$ 从点 $\bar{x}$ 出发沿方向 d 在 $\bar{x}$ 附近是下降的；当 $\dfrac{\partial f(\bar{x})}{\partial d}>0$ 时，$f(x)$ 从点 $\bar{x}$ 出发沿方向 d 在 $\bar{x}$ 附近是上升的. 方向导数的正负符号决定了函数的升降，升降快慢由它的绝对值大小决定. 绝对值越大，升降的速度就越快. 所以，方向导数 $\dfrac{\partial f(\bar{x})}{\partial d}$ 又可以称为函数 $f(x)$ 在点 $\bar{x}$ 处沿方向 d 的变化率. 根据 Cauchy 不等式，有

$$\left|\frac{\partial f(\bar{x})}{\partial d}\right|=|\nabla f(\bar{x})^{\mathrm{T}}e|\leqslant\|\nabla f(\bar{x})\|\cdot\|e\|=\|\nabla f(\bar{x})\|,$$

且当 $e=\dfrac{\nabla f(\bar{x})}{\|\nabla f(\bar{x})\|}$ 时，$\left|\dfrac{\partial f(\bar{x})}{\partial d}\right|$ 取得最大值. 由此可知，梯度方向是函数值上升最快的方向，而函数值下降最快的方向是负梯度方向. 因此，我们把负梯度方向叫做最速下降方向(steepest descent direction).

根据定理 1.2.2，我们还能得到下述重要结论.

定理 1.2.3 设 $f:\mathbb{R}^n\to\mathbb{R}$，$\bar{x}\in\mathbb{R}^n$，且 $f(x)$ 在点 $\bar{x}$ 处可微. 如果存在非零向量 $d\in\mathbb{R}^n$，使得 $\nabla f(\bar{x})^{\mathrm{T}}d<0$，则 d 是 f 在 $\bar{x}$ 处的下降方向；而当 $\nabla f(\bar{x})^{\mathrm{T}}d>0$ 时，d 是 f 在 $\bar{x}$ 处的上升方向.

这个定理说明，与 f 在 $\bar{x}$ 处的梯度方向交成锐角的任何方向都是 f 在 $\bar{x}$ 处的上升方向；相反，与 f 在 $\bar{x}$ 处的梯度方向交成钝角的任何方向都是 f 在 $\bar{x}$ 处的下降方向.

下面，我们来看函数在一点处梯度的一个十分重要的几何性质.

设 $f:\mathbb{R}^n\to\mathbb{R}$，$\bar{x}\in\mathbb{R}^n$，则 $f(x)$ 过点 $\bar{x}$ 的等值面方程为

$$f(x)=f(\bar{x}).$$

记 $c=f(\bar{\boldsymbol{x}})$，$\boldsymbol{x}=(x_1,x_2,\cdots,x_n)^{\mathrm{T}}$，则上式即为

$$f(x_1,x_2,\cdots,x_n)=c. \tag{1.2.9}$$

今设 l 为该等值面上过点 $\bar{\boldsymbol{x}}=(\bar{x}_1,\bar{x}_2,\cdots,\bar{x}_n)^{\mathrm{T}}$ 的任一光滑曲线，则 l 可由如下参数形式表示：

$$l:\begin{cases} x_1=x_1(t), \\ x_2=x_2(t), \\ \cdots\cdots\cdots\cdots \\ x_n=x_n(t), \end{cases} \tag{1.2.10}$$

并且对应地有 $\bar{t}$ 使

$$\bar{x}_i=x_i(\bar{t}),i=1,2,\cdots,n.$$

显然，曲线 l 在点 $\bar{\boldsymbol{x}}$ 处的切向量为

$$\boldsymbol{T}(\bar{\boldsymbol{x}})=(x_1'(\bar{t}),x_2'(\bar{t}),\cdots,x_n'(\bar{t}))^{\mathrm{T}}.$$

将(1.2.10)式代入(1.2.9)式，得

$$f(x_1(t),x_2(t),\cdots,x_n(t))=c,$$

把上式两边对 t 求导，同时代入点 $\bar{t}$，得

$$\frac{\partial f(\bar{\boldsymbol{x}})}{\partial x_1}x_1'(\bar{t})+\frac{\partial f(\bar{\boldsymbol{x}})}{\partial x_2}x_2'(\bar{t})+\cdots+\frac{\partial f(\bar{\boldsymbol{x}})}{\partial x_n}x_n'(\bar{t})=0.$$

于是得到

$$\nabla f(\bar{\boldsymbol{x}})^{\mathrm{T}}\boldsymbol{T}(\bar{\boldsymbol{x}})=0.$$

上式说明，$f(\boldsymbol{x})$ 在点 $\bar{\boldsymbol{x}}$ 处的梯度与 $f(\boldsymbol{x})$ 过点 $\bar{\boldsymbol{x}}$ 处等值面上任一曲线 l 在该点的切线垂直，故而与过该点的切平面垂直. 或者说，$\nabla f(\bar{\boldsymbol{x}})$ 是曲面 $f(\boldsymbol{x})=f(\bar{\boldsymbol{x}})$ 在点 $\bar{\boldsymbol{x}}$ 处的一个法线方向向量(如图 1.2.1 所示).

我们再给出 n 元函数的二阶导数即 Hesse 矩阵的概念.

定义 1.2.5 设 $f:\mathbb{R}^n\to\mathbb{R}$，$\bar{\boldsymbol{x}}\in\mathbb{R}^n$. 如果 f 在点 $\bar{\boldsymbol{x}}$ 处对于自变量 $\boldsymbol{x}=(x_1,x_2,\cdots,x_n)^{\mathrm{T}}$ 的各分量的二阶偏导数

$$\frac{\partial^2 f(\bar{\boldsymbol{x}})}{\partial x_i\partial x_j}(i,j=1,2,\cdots,n)$$

都存在，则称函数 $f(\boldsymbol{x})$ 在点 $\bar{\boldsymbol{x}}$ 处二阶可导，并称矩阵

$$\nabla^2 f(\bar{\boldsymbol{x}})=\begin{bmatrix} \dfrac{\partial^2 f(\bar{\boldsymbol{x}})}{\partial x_1^2} & \dfrac{\partial^2 f(\bar{\boldsymbol{x}})}{\partial x_1\partial x_2} & \cdots & \dfrac{\partial^2 f(\bar{\boldsymbol{x}})}{\partial x_1\partial x_n} \\ \dfrac{\partial^2 f(\bar{\boldsymbol{x}})}{\partial x_2\partial x_1} & \dfrac{\partial^2 f(\bar{\boldsymbol{x}})}{\partial x_2^2} & \cdots & \dfrac{\partial^2 f(\bar{\boldsymbol{x}})}{\partial x_2\partial x_n} \\ \vdots & \vdots & & \vdots \\ \dfrac{\partial^2 f(\bar{\boldsymbol{x}})}{\partial x_n\partial x_1} & \dfrac{\partial^2 f(\bar{\boldsymbol{x}})}{\partial x_n\partial x_2} & \cdots & \dfrac{\partial^2 f(\bar{\boldsymbol{x}})}{\partial x_n^2} \end{bmatrix}$$

为 $f(\boldsymbol{x})$ 在点 $\bar{\boldsymbol{x}}$ 处的二阶导数或 Hesse 矩阵.

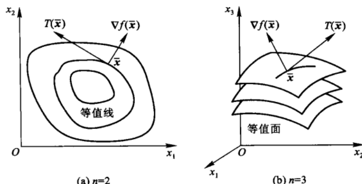


图 1.2.1 $\nabla f(\bar{\boldsymbol{x}})$ 与 $\boldsymbol{T}(\bar{\boldsymbol{x}})$ 的关系

在微积分中已经知道,当 $f(\boldsymbol{x})$ 在点 $\bar{\boldsymbol{x}}$ 处的所有二阶偏导数连续时,有

$$\frac{\partial^2 f(\bar{\boldsymbol{x}})}{\partial x_i \partial x_j}=\frac{\partial^2 f(\bar{\boldsymbol{x}})}{\partial x_j \partial x_i}, i,j=1,2,\cdots,n,$$

这时相应的 Hesse 矩阵为对称矩阵.

下面介绍在今后的计算中要用到的向量函数的导数.

定义 1.2.6 设 $\boldsymbol{h}:\mathbb{R}^n\to\mathbb{R}^m, \bar{\boldsymbol{x}}\in\mathbb{R}^n$,记

$$\boldsymbol{h}(\boldsymbol{x})=(h_1(\boldsymbol{x}),h_2(\boldsymbol{x}),\cdots,h_m(\boldsymbol{x}))^{\mathrm{T}},$$

如果 $h_i(\boldsymbol{x})(i=1,2,\cdots,m)$ 在点 $\bar{\boldsymbol{x}}$ 处对于自变量 $\boldsymbol{x}=(x_1,x_2,\cdots,x_n)^{\mathrm{T}}$ 的各分量的偏导数 $\dfrac{\partial h_i(\bar{\boldsymbol{x}})}{\partial x_j}(j=1,2,\cdots,n)$ 都存在,则称向量函数 $\boldsymbol{h}$ 在点 $\bar{\boldsymbol{x}}$ 处是一阶可导的,并且称矩阵

$$\nabla_{m\times n}\boldsymbol{h}(\bar{\boldsymbol{x}})=\begin{bmatrix}\dfrac{\partial h_1(\bar{\boldsymbol{x}})}{\partial x_1} & \dfrac{\partial h_1(\bar{\boldsymbol{x}})}{\partial x_2} & \cdots & \dfrac{\partial h_1(\bar{\boldsymbol{x}})}{\partial x_n}\\ \dfrac{\partial h_2(\bar{\boldsymbol{x}})}{\partial x_1} & \dfrac{\partial h_2(\bar{\boldsymbol{x}})}{\partial x_2} & \cdots & \dfrac{\partial h_2(\bar{\boldsymbol{x}})}{\partial x_n}\\ \vdots & \vdots & & \vdots\\ \dfrac{\partial h_m(\bar{\boldsymbol{x}})}{\partial x_1} & \dfrac{\partial h_m(\bar{\boldsymbol{x}})}{\partial x_2} & \cdots & \dfrac{\partial h_m(\bar{\boldsymbol{x}})}{\partial x_n}\end{bmatrix}$$

为 $\boldsymbol{h}(\boldsymbol{x})$ 在点 $\bar{\boldsymbol{x}}$ 处的一阶导数或 Jacobi 矩阵,简记为 $\nabla\boldsymbol{h}(\bar{\boldsymbol{x}})$.

我们知道,n 元函数 $f:\mathbb{R}^n\to\mathbb{R}$ 的梯度是向量函数

$$\nabla f(\boldsymbol{x})=\left(\frac{\partial f(\boldsymbol{x})}{\partial x_1},\frac{\partial f(\boldsymbol{x})}{\partial x_2},\cdots,\frac{\partial f(\boldsymbol{x})}{\partial x_n}\right)^{\mathrm{T}}.$$

由定义 1.2.6 知,$\nabla f(\boldsymbol{x})$ 的一阶导数或 Jacobi 矩阵为

$$\nabla_{n\times n}\nabla f(\boldsymbol{x})=\begin{bmatrix}\frac{\partial}{\partial x_1}\left(\frac{\partial f(\boldsymbol{x})}{\partial x_1}\right) & \frac{\partial}{\partial x_2}\left(\frac{\partial f(\boldsymbol{x})}{\partial x_1}\right) & \cdots & \frac{\partial}{\partial x_n}\left(\frac{\partial f(\boldsymbol{x})}{\partial x_1}\right)\\ \frac{\partial}{\partial x_1}\left(\frac{\partial f(\boldsymbol{x})}{\partial x_2}\right) & \frac{\partial}{\partial x_2}\left(\frac{\partial f(\boldsymbol{x})}{\partial x_2}\right) & \cdots & \frac{\partial}{\partial x_n}\left(\frac{\partial f(\boldsymbol{x})}{\partial x_2}\right)\\ \vdots & \vdots & & \vdots\\ \frac{\partial}{\partial x_1}\left(\frac{\partial f(\boldsymbol{x})}{\partial x_n}\right) & \frac{\partial}{\partial x_2}\left(\frac{\partial f(\boldsymbol{x})}{\partial x_n}\right) & \cdots & \frac{\partial}{\partial x_n}\left(\frac{\partial f(\boldsymbol{x})}{\partial x_n}\right)\end{bmatrix}$$

$$=\begin{bmatrix}\frac{\partial^2 f(\boldsymbol{x})}{\partial x_1^2} & \frac{\partial^2 f(\boldsymbol{x})}{\partial x_1\partial x_2} & \cdots & \frac{\partial^2 f(\boldsymbol{x})}{\partial x_1\partial x_n}\\ \frac{\partial^2 f(\boldsymbol{x})}{\partial x_2\partial x_1} & \frac{\partial^2 f(\boldsymbol{x})}{\partial x_2^2} & \cdots & \frac{\partial^2 f(\boldsymbol{x})}{\partial x_2\partial x_n}\\ \vdots & \vdots & & \vdots\\ \frac{\partial^2 f(\boldsymbol{x})}{\partial x_n\partial x_1} & \frac{\partial^2 f(\boldsymbol{x})}{\partial x_n\partial x_2} & \cdots & \frac{\partial^2 f(\boldsymbol{x})}{\partial x_n^2}\end{bmatrix}=\nabla^2 f(\boldsymbol{x}).$$

据此可知,函数梯度的 Jocobi 矩阵即为该函数的 Hesse 矩阵.

例 1.2.1 设 $\boldsymbol{a}\in\mathbb{R}^n,\boldsymbol{x}\in\mathbb{R}^n,b\in\mathbb{R}$,求线性函数

$$f(\boldsymbol{x})=\boldsymbol{a}^{\mathrm{T}}\boldsymbol{x}+b \tag{1.2.11}$$

在任意点 $\boldsymbol{x}$ 处的梯度和 Hesse 矩阵.

解 设 $\boldsymbol{a}=(a_1,a_2,\cdots,a_n)^{\mathrm{T}},\boldsymbol{x}=(x_1,x_2,\cdots,x_n)^{\mathrm{T}}$,则(1.2.11)式即为

$$f(x_1,x_2,\cdots,x_n)=\sum_{k=1}^{n}a_kx_k+b.$$

因 $\frac{\partial f(\boldsymbol{x})}{\partial x_k}=a_k(k=1,2,\cdots,n)$,故得 $\nabla f(\boldsymbol{x})=\boldsymbol{a}$. 又因

$$\frac{\partial^2 f(\boldsymbol{x})}{\partial x_i\partial x_j}=0(i,j=1,2,\cdots,n),$$

则 $f(\boldsymbol{x})$ 的 Hesse 矩阵为 n 阶零矩阵,即 $\nabla^2 f(\boldsymbol{x})=\boldsymbol{0}$. □

例 1.2.2 设 $\boldsymbol{Q}\in\mathbb{R}^{n\times n}$ 是对称矩阵,$\boldsymbol{b}\in\mathbb{R}^n,c\in\mathbb{R}$,二次函数

$$f(\boldsymbol{x})=\frac{1}{2}\boldsymbol{x}^{\mathrm{T}}\boldsymbol{Q}\boldsymbol{x}+\boldsymbol{b}^{\mathrm{T}}\boldsymbol{x}+c. \tag{1.2.12}$$

求 $f(\boldsymbol{x})$ 在任意点 $\boldsymbol{x}$ 处的梯度和 Hesse 矩阵.

解 设

$$\boldsymbol{Q}=(q_{ij})_{n\times n},\boldsymbol{x}=(x_1,x_2,\cdots,x_n)^{\mathrm{T}},\boldsymbol{b}=(b_1,b_2,\cdots,b_n)^{\mathrm{T}},$$

则(1.2.12)式即为

$$f(x_1,x_2,\cdots,x_n)=\frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}q_{ij}x_ix_j+\sum_{k=1}^{n}b_kx_k+c.$$

将它对各变量 $x_i(i=1,2,\cdots,n)$ 求偏导数,有

$$\nabla f(\boldsymbol{x}) = \begin{bmatrix} \dfrac{\partial f(\boldsymbol{x})}{\partial x_1} \\ \dfrac{\partial f(\boldsymbol{x})}{\partial x_2} \\ \vdots \\ \dfrac{\partial f(\boldsymbol{x})}{\partial x_n} \end{bmatrix} = \begin{bmatrix} \sum_{j=1}^{n} q_{1j}x_j + b_1 \\ \sum_{j=1}^{n} q_{2j}x_j + b_2 \\ \vdots \\ \sum_{j=1}^{n} q_{nj}x_j + b_n \end{bmatrix} = \begin{bmatrix} \sum_{j=1}^{n} q_{1j}x_j \\ \sum_{j=1}^{n} q_{2j}x_j \\ \vdots \\ \sum_{j=1}^{n} q_{nj}x_j \end{bmatrix} + \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix},$$

因而得到

$$\nabla f(\boldsymbol{x}) = \boldsymbol{Q}\boldsymbol{x} + \boldsymbol{b}.$$

再对$\dfrac{\partial f(\boldsymbol{x})}{\partial x_i} = \sum_{j=1}^{n} q_{ij}x_j + b_i(i = 1,2,\cdots,n)$求偏导得到

$$\frac{\partial^2 f(\boldsymbol{x})}{\partial x_i \partial x_j} = q_{ij}, i,j = 1,2,\cdots,n.$$

于是

$$\nabla^2 f(\boldsymbol{x}) = \begin{bmatrix} q_{11} & q_{12} & \cdots & q_{1n} \\ q_{21} & q_{22} & \cdots & q_{2n} \\ \vdots & \vdots & & \vdots \\ q_{n1} & q_{n2} & \cdots & q_{nn} \end{bmatrix} = \boldsymbol{Q}.$$ □

例 1.2.3 设 $\varphi(t) = f(\boldsymbol{x} + t\boldsymbol{d})$，其中 $f:\mathbb{R}^n \to \mathbb{R}$ 二阶可导，$\boldsymbol{x} \in \mathbb{R}^n$，$\boldsymbol{d} \in \mathbb{R}^n$，$t \in \mathbb{R}$，试求 $\varphi'(t)$，$\varphi''(t)$.

解 由多元复合函数微分法知

$$\varphi'(t) = \nabla f(\boldsymbol{x} + t\boldsymbol{d})^{\mathrm{T}}\boldsymbol{d},$$
$$\varphi''(t) = \boldsymbol{d}^{\mathrm{T}} \nabla^2 f(\boldsymbol{x} + t\boldsymbol{d})\boldsymbol{d}.$$ □

最后，我们给出 n 元函数在一点处的 Taylor 展式，它在最优化的理论和方法研究中有十分重要的应用.

定理 1.2.4 设 $f:\mathbb{R}^n \to \mathbb{R}$，$\bar{\boldsymbol{x}} \in \mathbb{R}^n$，如果 $f(\boldsymbol{x})$ 在 $\bar{\boldsymbol{x}}$ 的某邻域内具有二阶连续偏导数，则 f 在点 $\bar{\boldsymbol{x}}$ 处有 Taylor 展式

$$f(\bar{\boldsymbol{x}} + \Delta\boldsymbol{x}) = f(\bar{\boldsymbol{x}}) + \nabla f(\bar{\boldsymbol{x}})^{\mathrm{T}}\Delta\boldsymbol{x} + \frac{1}{2}\Delta\boldsymbol{x}^{\mathrm{T}} \nabla^2 f(\bar{\boldsymbol{x}} + \theta\Delta\boldsymbol{x})\Delta\boldsymbol{x}(0 < \theta < 1). \tag{1.2.13}$$

证明 设 $\varphi(t) = f(\bar{\boldsymbol{x}} + t\Delta\boldsymbol{x})$，$t \in [0,1]$，则 $\varphi(0) = f(\bar{\boldsymbol{x}})$，$\varphi(1) = f(\bar{\boldsymbol{x}} + \Delta\boldsymbol{x})$. 按一元函数 Taylor 公式将 $\varphi(t)$ 在 $t = 0$ 处展开有

$$\varphi(t) = \varphi(0) + \varphi'(0)t + \frac{1}{2}\varphi''(\theta)t^2 \quad (0 < \theta < 1). \tag{1.2.14}$$

从例 1.2.3 得知

$$\varphi'(0)=\nabla f(\bar{x})^{\mathrm{T}}\Delta x,\varphi''(\theta)=(\Delta x)^{\mathrm{T}}\nabla^2 f(\bar{x}+\theta\Delta x)\Delta x.$$

把它们代入(1.2.14)式,并令 $t=1$,即证得(1.2.13)式. □

由于 $\nabla^2 f(x)$ 的每一分量在点 $\bar{x}$ 处连续,故

$$\frac{\partial^2 f(\bar{x}+\theta\Delta x)}{\partial x_i\partial x_j}=\frac{\partial^2 f(\bar{x})}{\partial x_i\partial x_j}+\delta_{ij},i,j=1,2,\cdots,n,$$

其中 $\lim\limits_{\|\Delta x\|\to 0}\delta_{ij}=0$. 于是,(1.2.13)式也可写作

$$f(\bar{x}+\Delta x)=f(\bar{x})+\nabla f(\bar{x})^{\mathrm{T}}\Delta x+\frac{1}{2}\Delta x^{\mathrm{T}}\nabla^2 f(\bar{x})\Delta x+o(\|\Delta x\|^2). \tag{1.2.15}$$

我们分别把(1.2.7)式和(1.2.15)式叫做 $f(\bar{x})$ 在点 $\bar{x}$ 的一阶 Taylor 展式和二阶 Taylor 展式.

若记 $x=\bar{x}+\Delta x$,在(1.2.7)式和(1.2.15)式中略去高阶无穷小量的项后,相应地有近似关系式

$$f(x)\approx f(\bar{x})+\nabla f(\bar{x})^{\mathrm{T}}(x-\bar{x}), \tag{1.2.16}$$

$$f(x)\approx f(\bar{x})+\nabla f(\bar{x})^{\mathrm{T}}(x-\bar{x})+\frac{1}{2}(x-\bar{x})^{\mathrm{T}}\nabla^2 f(\bar{x})(x-\bar{x}). \tag{1.2.17}$$

通常,把(1.2.16)式的右边叫做函数 $f(x)$ 在点 $\bar{x}$ 处的线性逼近函数;把(1.2.17)式的右边叫做函数 $f(x)$ 在点 $\bar{x}$ 处的二次逼近函数.

1.3 最优化问题的图解法

为了讨论二维最优化问题的图解法,下面给出最优化问题(1.1.1)的最优解和极小点的正式定义.

定义 1.3.1 设 $f:\mathbb{R}^n\to\mathbb{R}$ 为目标函数,$S\subseteq\mathbb{R}^n$ 为可行域,$\bar{x}\in S$.

(1) 若对一切 $x\in S$,都有

$$f(x)\geqslant f(\bar{x}),$$

则称 $\bar{x}$ 为 $f(x)$ 在 S 上的全局(或整体)极小点(global minimum point),或者说,$\bar{x}$ 是最小化问题(1.1.1)的全局(或整体)最优解(global optimal solution),并称 $f(\bar{x})$ 为问题(1.1.1)的最优值(optimal value).

(2) 若对一切 $x\in S$ 但 $x\neq\bar{x}$,都有

$$f(x)>f(\bar{x}),$$

则称 $\bar{x}$ 为 $f(x)$ 在 S 上的严格全局(或整体)极小点(strictly globalminimum point).

(3) 若存在 $\bar{x}$ 的 δ 邻域

$$N_\delta(\bar{x}) = \{x \in \mathbb{R}^n \mid \|x - \bar{x}\| < \delta\}, \delta > 0,$$
使得对一切 $x \in N_\delta(\bar{x}) \cap S$，都有
$$f(x) \geqslant f(\bar{x}),$$
则称 $\bar{x}$ 为 $f(x)$ 在 S 上的局部极小点(local minimum point)，或者说，$\bar{x}$ 为极小化问题(1.1.1)的局部最优解(local optimal solution).

(4) 若存在 $\bar{x}$ 的 δ 邻域 $N_\delta(\bar{x})$ ($\delta > 0$)，使得对一切 $x \in N_\delta(\bar{x}) \cap S$ 但 $x \neq \bar{x}$，都有
$$f(x) > f(\bar{x}).$$
则称 $\bar{x}$ 为 $f(x)$ 在 S 上的严格局部极小点(strictly local minimum point).

由以上最优解的定义可以看出，$\bar{x}$ 是局部极小点是指在 S 的以 $\bar{x}$ 为中心的某个邻域内，$f(x)$ 在 $\bar{x}$ 处取到最小值. 而 $\bar{x}$ 是全局极小点是指在整个可行域 S 上，$f(x)$ 在点 $\bar{x}$ 处取得最小值. 局部极小点不一定是全局极小点，但全局极小点一定是局部极小点.

实际问题通常是求全局极小点，但最优化中绝大多数方法是求局部极小点. 在实际中，希望一个局部极小点是全局的，可以用一系列的初始点运用算法得出若干极小点，然后通过比较加以检验. 一般来说，即使只找到一个局部解也不能保证这个解就是全局最优解. 幸运的是，从实际中导出的问题，目标函数通常是具有单个极值的"良性"函数，所以采用数值方法得出的最优化问题的局部最优解，从实用的角度来说，并没有多大缺点.

如果考虑数值计算的精度，可以用
$$f(x) - r \geqslant f(\bar{x})$$
代替 $f(x) \geqslant f(\bar{x})$，其中 r 是某个很小的正数.

最后，我们来介绍二维问题的图解法. 从下面几个具体例子中，进一步加深对最优化基本概念的理解.

例 1.3.1 求解无约束极小化问题
$$\min (x_1 - 2)^2 + (x_2 - 1)^2. \tag{1.3.1}$$

这是定义在 Ox_1x_2 平面 $\mathbb{R}^2$ 上的无约束极小化问题，目标函数 $f(x) = (x_1 - 2)^2 + (x_2 - 1)^2$ 在 Ox_1x_2f 三维空间中代表一张曲面 S(这里为旋转抛物面). 从图 1.3.1 中的几何图形可以看出，$f(x)$ 在顶点(2,1)处达到最小值 0. 然而，通常的作法是用平行于 Ox_1x_2 面的平面截曲面 S，将其截痕曲线投影到 Ox_1x_2 面上得到曲面 S 的等值线图. 在同一等值线上，目标函数 $f(x)$ 的值相同. 求解极小化问题，无非是在可行域中找一点，使该点处的等值线(等值面)具有最小值. 由图知，过 $\bar{x} = (2,1)^T$ 的等值线具有最小值 0，所以 $\bar{x} = (2,1)^T$ 为问题(1.3.1)的全局极小点，最优值为 0. □

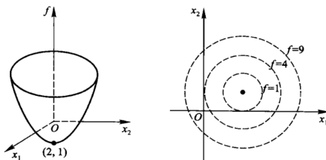


图 1.3.1 例 1.3.1 的图解

例 1.3.2 用图解法求解约束极小化问题

$$\begin{cases} \min & (x_1-2)^2+(x_2-1)^2; \\ \text{s.t.} & x_1+x_2-5=0. \end{cases} \tag{1.3.2}$$

从图 1.3.2 可以看出，最优解在等值线与约束直线的切点 $\bar{x}=(3,2)^{\mathrm{T}}$ 处取得. $\bar{x}$ 为问题(1.3.2)的全局最优解，最优值为 $f(\bar{x})=2$. □

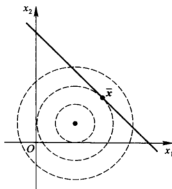


图 1.3.2 例 1.3.2 的图解

例 1.3.3 用图解法解下述约束极小化问题

$$\begin{cases} \min & (x_1-2)^2+(x_2-1)^2; \\ \text{s.t.} & x_1+x_2^2-5x_2=0, \\ & x_1+x_2-5\geqslant 0, \\ & x_1,x_2\geqslant 0. \end{cases} \tag{1.3.3}$$

解 (1) 画出目标函数的等值线.

令

$$f(\boldsymbol{x})=(x_1-2)^2+(x_2-1)^2=c(c\text{ 为非负常数}),$$

在几何上它们表示一族中心在$(2,1)^{\mathrm{T}}$的同心圆.

(2) 画出约束区域.

首先画出等式约束 $x_1+x_2^2-5x_2=0$ 的图形,是一条抛物线,再画出不等式约束所代表的区域.通常先画不等式约束的边界曲线,然后根据不等式的符号定出约束区域应在该边界曲线的哪一侧.

在图 1.3.3 中,抛物线弧 $ABCD$ 为约束集.由图可知,当动点从 A 出发沿抛物线弧 $ABCD$ 移动时,在弧 AB 段,目标函数下降;在弧 BC 段,目标函数上升;在弧 CD 段,目标函数值又下降.易知,D 是可行域上使目标函数值最小的点,它是全局最优点,其坐标由方程组

$$\begin{cases}x_1+x_2^2-5x_2=0,\\x_1+x_2-5=0\end{cases}$$

确定.从而求得问题(1.3.3)的全局最优点 $\bar{\boldsymbol{x}}=(4,1)^{\mathrm{T}}$(该方程组的另一解为点 $A(0,5)^{\mathrm{T}}$,不是最优点).

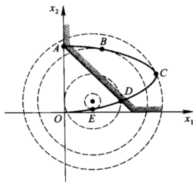


图 1.3.3　例 1.3.3 的图解

从几何上,我们还可观察到约束曲线弧 $ABCD$ 与等值线共有两个切点 B 和 C,这里 B、C 分别是问题(1.3.3)的局部极小点和局部极大点.抛物线与等值线的另一切点 E 不在问题(1.3.3)的可行域中. □

从以上几个例子可以看出,二维最优化问题具有鲜明的几何解释,因而可以借助于图解法求得其最优解.我们可以象征性地把这种几何解释推广到 n 维空间,这对于理解和掌握有关理论和方法是大有帮助的.

习 题 一

1. 试建立如下投资问题的数学模型：假设某公司在某一计划期内可用于投资的总金额为 b 万元，可供选择的投资项目共有 n 个. 已知第 i 个项目的投资额为 a_i 万元，收益额为 c_i 万元. 问如何进行投资，才能使利润率最大？

2. 计划在一段起伏的地形上修筑一条公路. 假设建筑的费用与增添或清除的泥土总量成正比. 设 T 为公路之长. 并设 $c(t)$ 是在任一给定 $t\in[0,T]$ 处地形的已知高度. 为避免公路过分倾斜，最大坡度不应当超过 b_1；同时，为减少公路的崎岖，要求公路坡度的变化率不超过 b_2. 试寻求描述公路高度 $y(t)$ $(t\in[0,T])$ 的方程.

3. 给定空间中 N 个点 $M_i(x_i,y_i,z_i)$ $(i=1,2,\cdots,N)$，问如何找一个最小的球完全包含这 N 个点，并使球心位于中心为原点的单位球面的第一卦限上？试写出该问题的最优化模型.

4. 用图解法求解下列问题：

$$(1)\begin{cases}\min & x_1^2+x_2;\\ \text{s.t.} & -x_1-x_2+1\geqslant 0,\\ & x_1^2+x_2^2-9=0.\end{cases}$$

$$(2)\begin{cases}\min & x_1^2+(x_2-1)^2+1;\\ \text{s.t.} & x_1-x_2^2+2\geqslant 0,\\ & x_2\leqslant x_1,\\ & x_1\geqslant 0,\\ & x_2\geqslant 0.\end{cases}$$

5. 求下列函数的梯度和 Hesse 矩阵：

(1) $f(x_1,x_2,x_3)=x_1\mathrm{e}^{x_1+x_2+x_3}$；

(2) $f(\boldsymbol{x})=\boldsymbol{g}(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{g}(\boldsymbol{x})$，其中 $\boldsymbol{g}:\mathbb{R}^n\to\mathbb{R}^m$.

6. 设 $\boldsymbol{x}=(x_1,x_2)^{\mathrm{T}}$，计算 Powell 函数

$$f(\boldsymbol{x})=x_1^4+x_1x_2+(1+x_2)^2$$

的梯度和 Hesse 矩阵，并求 $\nabla f(\boldsymbol{0})$，$\nabla^2 f(\boldsymbol{0})$，由此验证 $\nabla^2 f(\boldsymbol{0})$ 是非正定的.

7. 设 $\boldsymbol{x}=(x_1,x_2)^{\mathrm{T}}$，对 Rosenbrock 函数

$$f(\boldsymbol{x})=100(x_2-x_1^2)^2+(1-x_1)^2,$$

写出它的梯度和 Hesse 矩阵，并且验证对于满足 $\nabla f(\bar{\boldsymbol{x}})=\boldsymbol{0}$ 的 $\bar{\boldsymbol{x}}$，$\nabla^2 f(\bar{\boldsymbol{x}})$ 是正定的.

8. 对于正定二次函数

$$f(\boldsymbol{x})=\frac{1}{2}\boldsymbol{x}^{\mathrm{T}}\boldsymbol{Q}\boldsymbol{x}+\boldsymbol{b}^{\mathrm{T}}\boldsymbol{x}+c,$$

其中 $\boldsymbol{Q}$ 为正定矩阵，证明 $\bar{\boldsymbol{x}}=-\boldsymbol{Q}^{-1}\boldsymbol{b}$ 为其严格全局极小点.

9. 设方向 $\boldsymbol{d}$ 是函数 f 在点 $\boldsymbol{x}$ 处的下降方向，令

$$\boldsymbol{H}=\boldsymbol{I}-\frac{\boldsymbol{d}\boldsymbol{d}^{\mathrm{T}}}{\boldsymbol{d}^{\mathrm{T}}\nabla f(\boldsymbol{x})}-\frac{\nabla f(\boldsymbol{x})\nabla f(\boldsymbol{x})^{\mathrm{T}}}{\nabla f(\boldsymbol{x})^{\mathrm{T}}\nabla f(\boldsymbol{x})},$$

其中 $\boldsymbol{I}$ 为单位矩阵，证明方向 $\boldsymbol{p}=-\boldsymbol{H}\nabla f(\boldsymbol{x})$ 也是函数 f 在点 $\boldsymbol{x}$ 处的下降方向.

第二章 凸性

凸性是最优化理论必须涉及的基本概念. 具有凸性的非线性规划模型是一类重要的特殊模型,它在最优化的理论证明及算法研究中具有非常重要的作用.

本章先介绍凸集、凸集分离定理和多胞形的表示定理,再阐述凸函数的概念及性质,最后讨论具有凸性的非线性规划——凸规划问题.

2.1 凸 集

凸集是最优化的基础. 下面我们要介绍凸集的定义和几个基本性质.

2.1.1 凸集的概念及性质

首先给出$\mathbb{R}^n$中有限个点的凸组合的概念.

给定 m 个点 $\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_m\in\mathbb{R}^n$ 和实数 $\lambda_1,\lambda_2,\cdots,\lambda_m$,称表达式

$$\lambda_1\boldsymbol{x}_1+\lambda_2\boldsymbol{x}_2+\cdots+\lambda_m\boldsymbol{x}_m \tag{2.1.1}$$

为点 $\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_m$ 的线性组合(linear combination).

特别地,当 $\lambda_1+\lambda_2+\cdots+\lambda_m=1$ 时,称(2.1.1)式为点 $\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_m$ 的仿射组合(affine combination);

当 $\lambda_1+\lambda_2+\cdots+\lambda_m=1$ 且 $\lambda_1,\lambda_2,\cdots,\lambda_m\geqslant 0$ 时,称(2.1.1)式为点 $\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_m$ 的凸组合(convex combination);

当 $\lambda_1+\lambda_2+\cdots+\lambda_m=1$ 且 $\lambda_1,\lambda_2,\cdots,\lambda_m>0$ 时,称(2.1.1)式为点 $\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_m$ 的严格凸组合(strictly convex combination);

当 $\lambda_1,\lambda_2,\cdots,\lambda_m\geqslant 0$ 时,称(2.1.1)式为点 $\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_m$ 的凸锥组合(convex cone combination).

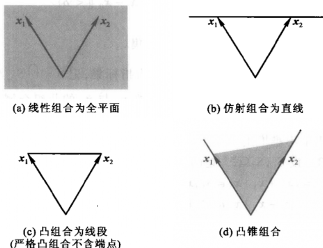
图 2.1.1 给出了二维情况下两点 $\boldsymbol{x}_1,\boldsymbol{x}_2$ 的组合图示. 从图 2.1.1(c)可以看出,两点 $\boldsymbol{x}_1,\boldsymbol{x}_2$ 的凸组合在几何上表示连接这两点的线段,记作$\overline{\boldsymbol{x}_1\boldsymbol{x}_2}$,即

$$\overline{\boldsymbol{x}_1\boldsymbol{x}_2}=\{\boldsymbol{x}\in\mathbb{R}^n\mid \boldsymbol{x}=\lambda\boldsymbol{x}_1+(1-\lambda)\boldsymbol{x}_2,\lambda\in[0,1]\}.$$

现在给出凸集的定义.

定义 2.1.1 设 $S\subseteq\mathbb{R}^n$,如果对任意 $\boldsymbol{x}_1,\boldsymbol{x}_2\in S$,有

$$\lambda\boldsymbol{x}_1+(1-\lambda)\boldsymbol{x}_2\in S,\ \forall\lambda\in[0,1],$$



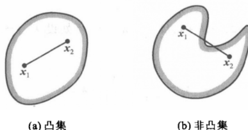
(a) 线性组合为全平面　　(b) 仿射组合为直线

(c) 凸组合为线段
(严格凸组合不含端点)　　(d) 凸锥组合

图 2.1.1　二维情况下两点的组合

则称 S 为$\mathbb{R}^n$ 中的凸集(convex set).

从定义可以看出,凸集是这样的集合,连接其中任意两点的线段上所有的点都属于此集合. 图 2.1.2 画出了凸集和非凸集的图形.



(a) 凸集　　(b) 非凸集

图 2.1.2　凸集与非凸集

由定义 2.1.1,容易验证下列结论是正确的:

(1) 空集$\varnothing$和全空间$\mathbb{R}^n$ 是凸集.

(2) 设 $\boldsymbol{a}\in\mathbb{R}^n,\boldsymbol{a}\neq\boldsymbol{0}$,且 $\alpha\in\mathbb{R}$,则超平面(hyperplane)

$$H=\{\boldsymbol{x}\in\mathbb{R}^n\mid\boldsymbol{a}^{\mathrm{T}}\boldsymbol{x}=\alpha\}$$

是凸集;半空间(halfspace)

$$H^-=\{\boldsymbol{x}\in\mathbb{R}^n\mid\boldsymbol{a}^{\mathrm{T}}\boldsymbol{x}\leqslant\alpha\},H^+=\{\boldsymbol{x}\in\mathbb{R}^n\mid\boldsymbol{a}^{\mathrm{T}}\boldsymbol{x}\geqslant\alpha\}$$

也都是凸集.

(3) 设 $\boldsymbol{x}_0\in\mathbb{R}^n,\delta>0$, 则点 $\boldsymbol{x}_0$ 的 δ 邻域,即开超球(open hypershere)

$$N_\delta(\boldsymbol{x}_0) = \{\boldsymbol{x} \in \mathbb{R}^n \mid \|\boldsymbol{x} - \boldsymbol{x}_0\| < \delta\}$$

是凸集.

定理 2.1.1 任意多个凸集的交集是凸集.

证明 设 $S_i \subseteq \mathbb{R}^n (i \in I)$是凸集,其中 I 为指标集,记 $S = \bigcap_{i \in I} S_i$. 对于任意 $\boldsymbol{x}_1$, $\boldsymbol{x}_2 \in S$,有 $\boldsymbol{x}_1, \boldsymbol{x}_2 \in S_i (\forall i \in I)$. 因 S_i 为凸集,故 $\boldsymbol{x}_1$ 与 $\boldsymbol{x}_2$ 的凸组合属于 $S_i (i \in I)$,从而该凸组合也属于 S,即 S 为凸集. □

由定义 2.1.1 不难推出

定理 2.1.2 设 $S_1, S_2 \subseteq \mathbb{R}^n$ 是凸集,$\lambda \in \mathbb{R}$,则

(1) $S_1 + S_2 = \{\boldsymbol{x}_1 + \boldsymbol{x}_2 \mid \boldsymbol{x}_1 \in S_1, \boldsymbol{x}_2 \in S_2\}$ 是凸集;

(2) $S_1 - S_2 = \{\boldsymbol{x}_1 - \boldsymbol{x}_2 \mid \boldsymbol{x}_1 \in S_1, \boldsymbol{x}_2 \in S_2\}$ 是凸集;

(3) $\lambda S_1 = \{\lambda \boldsymbol{x} \mid \boldsymbol{x} \in S_1\}$ 是凸集. □

定理 2.1.3 $S \subseteq \mathbb{R}^n$ 是凸集当且仅当对于任意正整数 $m \geqslant 2$,S 中任意 m 个点 $\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_m$ 的一切凸组合都属于 S.

证明 充分性显然. 下证必要性. 设 S 是凸集,对 m 用数学归纳法证明. 当 $m = 2$ 时,由凸集的定义知结论成立. 假设 $m = k$ 时结论成立. 下面考虑$m = k + 1$时的情形. 令

$$\boldsymbol{x} = \sum_{i=1}^{k+1} \lambda_i \boldsymbol{x}_i,$$

其中 $\boldsymbol{x}_i \in S, \lambda_i \geqslant 0, i = 1, 2, \cdots, k+1$,且 $\sum_{i=1}^{k+1} \lambda_i = 1$. 不妨设 $\lambda_{k+1} \neq 1$(不然 $\boldsymbol{x} = \boldsymbol{x}_{k+1} \in S$,结论成立),记

$$\boldsymbol{y} = \sum_{i=1}^{k} \frac{\lambda_i}{1 - \lambda_{k+1}} \boldsymbol{x}_i,$$

有 $\boldsymbol{x} = (1 - \lambda_{k+1})\boldsymbol{y} + \lambda_{k+1}\boldsymbol{x}_{k+1}$. 又

$$\frac{\lambda_i}{1 - \lambda_{k+1}} \geqslant 0, i = 1, 2, \cdots, k; \sum_{i=1}^{k} \frac{\lambda_i}{1 - \lambda_{k+1}} = 1,$$

则由归纳假设知,$\boldsymbol{y} \in S$. 而 $\boldsymbol{x}_{k+1} \in S$ 且 S 是凸集,故 $\boldsymbol{x} \in S$. □

定义 2.1.2 设 $S \subseteq \mathbb{R}^n$,则 S 中任意有限个点的所有凸组合所构成的集合称为 S 的凸包(convex hull),记为 $H(S)$,即

$$H(S) = \left\{ \sum_{i=1}^{m} \lambda_i \boldsymbol{x}_i \,\middle|\, \boldsymbol{x}_i \in S, \lambda_i \geqslant 0, i = 1, 2, \cdots, m, \sum_{i=1}^{m} \lambda_i = 1, m \in \mathbb{N}_+ \right\},$$

其中$\mathbb{N}_+$为所有正整数的集合. 特别地,若 $S = \{\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_k\}$,则称 $H(S)$ 是由 $\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_k$ 所生成的凸包.

图 2.1.3 给出了一个集合的凸包的图形.

易知,集 S 的凸包 $H(S)$是凸集.

定理 2.1.4 设 $S \subseteq \mathbb{R}^n$,则 $H(S)$ 是 $\mathbb{R}^n$ 中所有包含 S 的凸集的交集.

证明 设 $\mathbb{R}^n$ 中所有包含 S 的凸集的交集为 C,则由定理 2.1.1 知,C 为凸集.

因为 $S \subseteq H(S)$,$H(S)$ 是凸集,所以 $C \subseteq H(S)$. 又注意到 $S \subseteq C$ 及 C 为凸集,而 $H(S)$ 是由 S 中的点即 C 中一部分点的凸组合构成的,因此由定理 2.1.3(必要性)可知 $H(S) \subseteq C$. 所以,$C = H(S)$. □

由定理 2.1.4 知,$H(S)$ 是包含集合 S 的最小凸集.

一类重要的凸集是凸锥.

定义 2.1.3 设 $S \subseteq \mathbb{R}^n$,$\boldsymbol{x}_0 \in S$,如果对一切 $\boldsymbol{x} \in S$ 及 $\lambda > 0$,有 $\boldsymbol{x}_0 + \lambda \boldsymbol{x} \in S$,则称 S 是以 $\boldsymbol{x}_0$ 为顶点(vertex)的锥(cone). 如果锥 S 又是凸集,则称 S 为凸锥(convex cone).

图 2.1.4 给出了凸锥和非凸锥的图例.

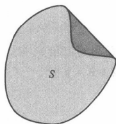
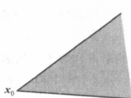
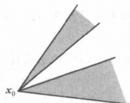


图 2.1.3 集合 S 的凸包 $H(S)$



(a) 凸锥



(b) 非凸锥

图 2.1.4 凸锥与非凸锥

设 $\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_m \in \mathbb{R}^n$,不难知道,集合

$$S = \left\{ \boldsymbol{x} \;\middle|\; \boldsymbol{x} = \sum_{i=1}^{m} \lambda_i \boldsymbol{x}_i, \lambda_i \geqslant 0, i = 1, 2, \cdots, m \right\}$$

是一个凸锥,我们称 S 为由 $\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_m$ 所生成的凸锥.

2.1.2 凸集分离定理与 Farkas 引理

凸集的许多重要应用都要涉及凸集的分离性质. 为此,下面讨论凸集的分离问题.

定义 2.1.4 设 $S_1, S_2 \subseteq \mathbb{R}^n$ 是非空集合,如果存在 $\boldsymbol{p} \in \mathbb{R}^n$,$\boldsymbol{p} \neq \boldsymbol{0}$ 及 $\alpha \in \mathbb{R}$,使

$$S_1 \subseteq H^- = \{\boldsymbol{x} \in \mathbb{R}^n \mid \boldsymbol{p}^{\mathrm{T}} \boldsymbol{x} \leqslant \alpha\},$$

$$S_2 \subseteq H^+ = \{\boldsymbol{x} \in \mathbb{R}^n \mid \boldsymbol{p}^{\mathrm{T}} \boldsymbol{x} \geqslant \alpha\},$$

则称超平面 $H = \{\boldsymbol{x} \in \mathbb{R}^n \mid \boldsymbol{p}^{\mathrm{T}} \boldsymbol{x} = \alpha\}$ 分离集合 S_1 和 S_2.

在介绍凸集分离定理之前,我们先给出闭凸集的一个性质.

定理 2.1.5 设 $S \subset \mathbb{R}^n$ 为非空闭凸集,$\boldsymbol{y} \in \mathbb{R}^n \setminus S$,则存在惟一的 $\bar{\boldsymbol{x}} \in S$,使得它

与 $\boldsymbol{y}$ 的距离为最小,即有

$$\|\bar{\boldsymbol{x}}-\boldsymbol{y}\| = \inf\{\|\boldsymbol{x}-\boldsymbol{y}\| \mid \boldsymbol{x}\in S\} > 0.$$

证明 考虑单位超球 $B=\{z\in\mathbb{R}^n \mid \|z\|\leqslant 1\}$,取充分大的 $\beta>0$,使

$$D = S\cap(\{\boldsymbol{y}\}+\beta B)\neq\varnothing.$$

因为 S 是闭集,$\{\boldsymbol{y}\}+\beta B$ 是有界闭集,所以 D 是非空有界闭集. 于是,连续函数 $f(x)=\|\boldsymbol{x}-\boldsymbol{y}\|$ 在 D 上取最小值. 设这个最小值在 $\bar{\boldsymbol{x}}\in D$ 处达到,即 $\bar{\boldsymbol{x}}$ 为 $\boldsymbol{y}$ 到 S 的最小距离点.

再证惟一性. 设存在 $\tilde{\boldsymbol{x}}\in S$,使

$$\|\bar{\boldsymbol{x}}-\boldsymbol{y}\| = \|\tilde{\boldsymbol{x}}-\boldsymbol{y}\| = r, \tag{2.1.2}$$

记 $\hat{\boldsymbol{x}}=\frac{1}{2}(\bar{\boldsymbol{x}}+\tilde{\boldsymbol{x}})$,则

$$\|\hat{\boldsymbol{x}}-\boldsymbol{y}\| \leqslant \frac{1}{2}\|\bar{\boldsymbol{x}}-\boldsymbol{y}\| + \frac{1}{2}\|\tilde{\boldsymbol{x}}-\boldsymbol{y}\| = r,$$

因为 S 是凸集,所以 $\hat{\boldsymbol{x}}\in S$. 由于 r 是最小距离,因此

$$\|\hat{\boldsymbol{x}}-\boldsymbol{y}\| = r. \tag{2.1.3}$$

根据平行四边形定律(对角线的平方和等于一组邻边平方和的二倍),有

$$\|\bar{\boldsymbol{x}}-\tilde{\boldsymbol{x}}\|^2 + 4\|\hat{\boldsymbol{x}}-\boldsymbol{y}\|^2 = 2\|\bar{\boldsymbol{x}}-\boldsymbol{y}\|^2 + 2\|\tilde{\boldsymbol{x}}-\boldsymbol{y}\|^2,$$

从而,由(2.1.2)和(2.1.3)两式得

$$\|\bar{\boldsymbol{x}}-\tilde{\boldsymbol{x}}\|^2 = 0,$$

所以 $\bar{\boldsymbol{x}}=\tilde{\boldsymbol{x}}$. □

现在给出凸集分离定理.

定理 2.1.6 设 $S\subset\mathbb{R}^n$ 是非空闭凸集,$\boldsymbol{y}\in\mathbb{R}^n\setminus S$,则存在 $\boldsymbol{p}\in\mathbb{R}^n$,$\boldsymbol{p}\neq\boldsymbol{0}$ 及 $\alpha\in\mathbb{R}$,使

$$\boldsymbol{p}^{\mathrm{T}}\boldsymbol{x}\leqslant\alpha<\boldsymbol{p}^{\mathrm{T}}\boldsymbol{y},\ \forall\,\boldsymbol{x}\in S,$$

即存在超平面 $H=\{\boldsymbol{x}\in\mathbb{R}^n \mid \boldsymbol{p}^{\mathrm{T}}\boldsymbol{x}=\alpha\}$ 分离 $\boldsymbol{y}$ 和 S.

证明 由于 $S\subset\mathbb{R}^n$ 为非空闭凸集,$\boldsymbol{y}\notin S$,因此由定理 2.1.5 知,存在 $\bar{\boldsymbol{x}}\in S$,使

$$\|\bar{\boldsymbol{x}}-\boldsymbol{y}\| = \inf\{\|\boldsymbol{x}-\boldsymbol{y}\| \mid \boldsymbol{x}\in S\} > 0.$$

因 S 为凸集,故对一切 $\boldsymbol{x}\in S$ 及 $\lambda\in(0,1)$,有

$$\lambda\boldsymbol{x}+(1-\lambda)\bar{\boldsymbol{x}}\in S.$$

于是

$$\begin{aligned}\|\bar{\boldsymbol{x}}-\boldsymbol{y}\|^2 &\leqslant \|\lambda\boldsymbol{x}+(1-\lambda)\bar{\boldsymbol{x}}-\boldsymbol{y}\|^2 = \|(\bar{\boldsymbol{x}}-\boldsymbol{y})+\lambda(\boldsymbol{x}-\bar{\boldsymbol{x}})\|^2\\ &= \|\bar{\boldsymbol{x}}-\boldsymbol{y}\|^2+\lambda^2\|\boldsymbol{x}-\bar{\boldsymbol{x}}\|^2+2\lambda(\bar{\boldsymbol{x}}-\boldsymbol{y})^{\mathrm{T}}(\boldsymbol{x}-\bar{\boldsymbol{x}}),\end{aligned}$$

从而

$$\lambda\|\boldsymbol{x}-\bar{\boldsymbol{x}}\|^2+2(\bar{\boldsymbol{x}}-\boldsymbol{y})^{\mathrm{T}}(\boldsymbol{x}-\bar{\boldsymbol{x}})\geqslant 0,\ \forall\,\lambda\in(0,1),\ \forall\,\boldsymbol{x}\in S.$$

在上式中令 $\lambda\to0^+$，得

$$(\bar{x}-y)^{T}(x-\bar{x})\geqslant 0,\ \forall x\in S.$$

记 $p=y-\bar{x}$，则 $p\neq 0$，且

$$p^{T}(x-\bar{x})\leqslant 0,\ \forall x\in S.$$

又记 $\alpha=p^{T}\bar{x}$，则有

$$p^{T}x\leqslant p^{T}\bar{x}=\alpha,\ \forall x\in S.$$

另一方面，因为

$$p^{T}y-\alpha=p^{T}(y-\bar{x})=\|y-\bar{x}\|^{2}>0,$$

所以 $p^{T}x\leqslant\alpha<p^{T}y,\ \forall x\in S$. □

作为凸集分离定理的应用，最后我们介绍在最优化理论中十分重要的Farkas引理.

定理 2.1.7（Farkas 引理） 设 $A\in\mathbb{R}^{m\times n}$，$b\in\mathbb{R}^{n}$，则下列两个关系式组有且仅有一组有解：

$$Ax\leqslant 0,b^{T}x>0; \tag{2.1.4}$$

$$A^{T}y=b,y\geqslant 0. \tag{2.1.5}$$

证明 设(2.1.5)式有解，即存在 $\bar{y}\geqslant 0$，使 $A^{T}\bar{y}=b$，若有 $\bar{x}$ 使 $A\bar{x}\leqslant 0$，则有

$$b^{T}\bar{x}=\bar{y}^{T}A\bar{x}\leqslant 0,$$

这表明(2.1.4)式无解.

再假设(2.1.5)式无解，记

$$S=\{z\mid z=A^{T}y,y\geqslant 0\},$$

则 S 是非空闭凸集，且 $b\notin S$，由定理 2.1.6，存在 $p\in\mathbb{R}^{n}$，$p\neq 0$ 及 $\alpha\in\mathbb{R}$，使

$$p^{T}z\leqslant\alpha<p^{T}b,\ \forall z\in S.$$

因 $0\in S$，故由上式知 $\alpha\geqslant 0$，从而 $p^{T}b>0$，于是

$$\alpha\geqslant p^{T}z=p^{T}A^{T}y=y^{T}Ap,\ \forall y\geqslant 0.$$

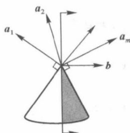
由于 $y\geqslant 0$，y 的分量可以任意大，因此 $Ap\leqslant 0$，这样就找到了 $p\in\mathbb{R}^{n}$，使 $Ap\leqslant 0$，$p^{T}b>0$，即(2.1.4)式有解. □

下面我们对 Farkas 引理作一几何解释. 设 A 的第 i 个行向量为 a_i，$i=1,2,\cdots,m$，则(2.1.4)式有解当且仅当凸锥 $\{x\mid Ax\leqslant 0\}$ 与半空间 $\{x\mid b^{T}x>0\}$ 的交不空，也就是说，(2.1.4)式有解当且仅当存在向量 x，它与各 a_i 的夹角为钝角或直角，而与 b 的夹角为锐角. (2.1.5)式有解当且仅当 b 在由 $a_1,a_2,\cdots,a_m$ 所生成的凸锥内. 图 2.1.5 能够帮助我们理解这两个问题之间的关系.

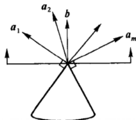
利用 Farkas 引理可推导下述的 Gordan 定理和择一性定理.

定理 2.1.8（Gordan 定理） 设 $A\in\mathbb{R}^{m\times n}$，则下列两个关系式组有且仅有一组有解：

$$Ax<0; \tag{2.1.6}$$



(2.1.4) 式有解(2.1.5) 式无解



(2.1.5) 式有解(2.1.4) 式无解

图 2.1.5 Farkas 引理的几何解释

$$A^{\mathrm{T}}y = \mathbf{0}, y \geqslant \mathbf{0}, y \neq \mathbf{0}. \tag{2.1.7}$$

证明 如果(2.1.6)式有解,即存在 $\bar{x} \in \mathbb{R}^n$,使 $A\bar{x} < \mathbf{0}$,则 $\forall \bar{y} \geqslant \mathbf{0}, \bar{y} \neq \mathbf{0}$,有 $\bar{y}^{\mathrm{T}} A\bar{x} < 0$,即 $\bar{x}^{\mathrm{T}} A^{\mathrm{T}} \bar{y} < 0$,这表明(2.1.7)式无解.

如果(2.1.6)式无解,则不存在 $\alpha < 0$ 及 $x \in \mathbb{R}^n$,使 $Ax \leqslant (\alpha, \alpha, \cdots, \alpha)^{\mathrm{T}}$. 记

$$\tilde{A} = (A, -e), \tilde{b} = (0, 0, \cdots, 0, -1)^{\mathrm{T}} \in \mathbb{R}^{n+1},$$

其中 $e = (1, 1, \cdots, 1)^{\mathrm{T}} \in \mathbb{R}^m$,于是不存在 $\alpha < 0$ 及 $x \in \mathbb{R}^n$,满足

$$\tilde{A}\begin{bmatrix} x \\ \alpha \end{bmatrix} \leqslant \mathbf{0}, \tilde{b}^{\mathrm{T}}\begin{bmatrix} x \\ \alpha \end{bmatrix} > 0,$$

即上式关系组无解. 于是由 Farkas 引理,下述关系组

$$\tilde{A}^{\mathrm{T}} y = \tilde{b}, y \geqslant \mathbf{0}$$

有解,即关系式组

$$A^{\mathrm{T}} y = \mathbf{0}, e^{\mathrm{T}} y = 1, y \geqslant \mathbf{0}$$

有解,这等价于(2.1.7)式有解. □

定理 2.1.9(择一性定理) 设 $A \in \mathbb{R}^{m \times n}, B \in \mathbb{R}^{p \times n}$,则关系式组

$$Ax < \mathbf{0}, Bx = \mathbf{0} \tag{2.1.8}$$

无解当且仅当存在 $u \in \mathbb{R}^m, u \geqslant \mathbf{0}, u \neq \mathbf{0}$ 和 $v \in \mathbb{R}^p$,满足

$$A^{\mathrm{T}} u + B^{\mathrm{T}} v = \mathbf{0}. \tag{2.1.9}$$

证明 式(2.1.8)无解等价于不存在 $\alpha < 0$ 及 $x \in \mathbb{R}^n$,满足

$$Ax \leqslant (\alpha, \alpha, \cdots, \alpha)^{\mathrm{T}}, Bx \leqslant \mathbf{0}, -Bx \leqslant \mathbf{0}.$$

记

$$\tilde{A} = \begin{bmatrix} A & -e \\ B & \mathbf{0} \\ -B & \mathbf{0} \end{bmatrix}, \tilde{b} = (0, 0, \cdots, 0, -1)^{\mathrm{T}} \in \mathbb{R}^{n+1},$$

其中 $e = (1, 1, \cdots, 1)^{\mathrm{T}} \in \mathbb{R}^m$,则不存在 $\alpha < 0$ 及 $x \in \mathbb{R}^n$,满足

$$\tilde{A}\begin{bmatrix} x \\ \alpha \end{bmatrix} \leqslant 0, \tilde{b}^{\mathrm{T}}\begin{bmatrix} x \\ \alpha \end{bmatrix} > 0. \tag{2.1.10}$$

这说明(2.1.8)式无解当且仅当(2.1.10)式无解.

根据 Farkas 引理,(2.1.10)式无解等价于关系式组

$$\tilde{A}^{\mathrm{T}} y = \tilde{b}, y \geqslant 0$$

有解. 记

$$y = \begin{bmatrix} u \\ w \\ z \end{bmatrix}, u \in \mathbb{R}^m, w \in \mathbb{R}^p, z \in \mathbb{R}^p,$$

则有

$$\begin{cases} A^{\mathrm{T}} u + B^{\mathrm{T}} w - B^{\mathrm{T}} z = 0, \\ e^{\mathrm{T}} u = 1, \\ u \geqslant 0, w \geqslant 0, z \geqslant 0. \end{cases}$$

由 $e^{\mathrm{T}} u = 1$ 知 $u \neq 0$,再令 $v = w - z$,则知(2.1.9)式成立. □

2.2 多胞形的表示定理

因为$\mathbb{R}^n$ 中的半空间是闭凸集,所以$\mathbb{R}^n$ 中任意个半空间的交集也是闭凸集. 我们称$\mathbb{R}^n$ 中有限个半空间的交集为多胞形(polytope),非空有界的多胞形称为多面体(polyhedron).

由于超平面

$$H = \{x \in \mathbb{R}^n \mid p^{\mathrm{T}} x = \alpha\}$$

可以表示为两个半空间

$$H^- = \{x \in \mathbb{R}^n \mid p^{\mathrm{T}} x \leqslant \alpha\}, H^+ = \{x \in \mathbb{R}^n \mid p^{\mathrm{T}} x \geqslant \alpha\}$$

的交集,即 $H = H^- \cap H^+$,因此超平面是多胞形.

设 $A \in \mathbb{R}^{m \times n}$,$b \in \mathbb{R}^m$,记

$$K = \{x \mid Ax = b, x \geqslant 0\}. \tag{2.2.1}$$

注意到 K 是 m 个超平面与 n 个半空间的交集,从而 K 是多胞形.

需要指出的是,本节中所出现的 K 均是(2.2.1)式中定义的多胞形,并且把矩阵 A 的第 j 列 p_j 称为 x 的第 j 个分量 x_j 对应的列向量.

定义 2.2.1 设 $S \subseteq \mathbb{R}^n$ 是凸集,$x_0 \in S$,若 x_0 不能表示成 S 中两个不同的点的严格凸组合,则称 x_0 是 S 的极点(extreme point).

换言之,x_0 为凸集 S 的极点是指:若存在 $x_1, x_2 \in S$,使

$$x_0 = \lambda_0 x_1 + (1 - \lambda_0) x_2, 0 < \lambda_0 < 1,$$

则必有 $x_1 = x_2$.

例如,在平面上,闭三角形区域的极点是三角形的 3 个顶点;闭圆域的圆周上任一点都是极点;开圆域没有极点;原点是每个象限区域的惟一极点;整个平面没有极点.

一般地,多胞形不一定有极点,但我们有

定理 2.2.1 若 $K\neq\varnothing$,则 K 必有极点.

证明 设 $\boldsymbol{x}_0\in K$,若 $\boldsymbol{x}_0$ 不是 K 的极点,则存在 $\bar{\boldsymbol{x}},\tilde{\boldsymbol{x}}\in K,\bar{\boldsymbol{x}}\neq\tilde{\boldsymbol{x}}$ 及 $\lambda_0\in(0,1)$,使

$$\boldsymbol{x}_0 = \lambda_0\bar{\boldsymbol{x}} + (1-\lambda_0)\tilde{\boldsymbol{x}}. \tag{2.2.2}$$

因 $\bar{\boldsymbol{x}}\neq\tilde{\boldsymbol{x}}$,故不妨设 $\tilde{\boldsymbol{x}}$ 至少有一个分量大于 $\bar{\boldsymbol{x}}$ 对应的分量,令

$$\lambda_1 = \min\left\{\frac{\tilde{x}_j}{\tilde{x}_j-\bar{x}_j}\,\middle|\,\tilde{x}_j>\bar{x}_j\right\},$$

则 $\lambda_1>0$,且

$$\boldsymbol{x}_1 = \lambda_1\bar{\boldsymbol{x}} + (1-\lambda_1)\tilde{\boldsymbol{x}} = \tilde{\boldsymbol{x}} + \lambda_1(\bar{\boldsymbol{x}}-\tilde{\boldsymbol{x}}) \geqslant \boldsymbol{0},$$

$$\boldsymbol{A}\boldsymbol{x}_1 = \lambda_1\boldsymbol{A}\bar{\boldsymbol{x}} + (1-\lambda_1)\boldsymbol{A}\tilde{\boldsymbol{x}} = \boldsymbol{b},$$

即 $\boldsymbol{x}_1\in K$. 又注意到,若 $\boldsymbol{x}_0$ 的某个分量为 0,则由(2.2.2)式知 $\bar{\boldsymbol{x}},\tilde{\boldsymbol{x}}$ 中与之对应的分量也为 0,从而 $\boldsymbol{x}_1$ 对应的分量也为 0,于是由 λ_1 的取法知,$\boldsymbol{x}_1$ 较 $\boldsymbol{x}_0$ 至少增加一个零分量. 如果 $\boldsymbol{x}_1$ 仍不是 K 的极点,则用 $\boldsymbol{x}_1$ 代替 $\boldsymbol{x}_0$ 重复上述做法,得到 $\boldsymbol{x}_2\in K$,且 $\boldsymbol{x}_2$ 比 $\boldsymbol{x}_1$ 至少增加一个零分量. 由于 K 中点的零分量个数不能无限地增加,因此,要么终止于某一步,找到 K 的一个极点 $\boldsymbol{x}_k$;要么找到 $\boldsymbol{x}_k=\boldsymbol{0}$,显然此时 $\boldsymbol{0}$ 为 K 的极点. □

下面的定理给出了 K 中极点的一个特征刻画.

定理 2.2.2 设 $\bar{\boldsymbol{x}}\in K$,则 $\bar{\boldsymbol{x}}$ 为 K 的极点的充要条件是 $\bar{\boldsymbol{x}}$ 的非零分量对应的列向量线性无关.

证明 设 $\bar{\boldsymbol{x}}\in K$,不失一般性,设 $\bar{\boldsymbol{x}}$ 的非零分量依次为 $\bar{x}_1,\bar{x}_2,\cdots,\bar{x}_r$,它们对应的列向量依次为 $\boldsymbol{p}_1,\boldsymbol{p}_2,\cdots,\boldsymbol{p}_r$,则

$$\boldsymbol{A}\bar{\boldsymbol{x}} = \bar{x}_1\boldsymbol{p}_1 + \bar{x}_2\boldsymbol{p}_2 + \cdots + \bar{x}_r\boldsymbol{p}_r = \boldsymbol{b}.$$

如果 $\boldsymbol{p}_1,\boldsymbol{p}_2,\cdots,\boldsymbol{p}_r$ 线性相关,则必有不全为 0 的数 $d_1,d_2,\cdots,d_r$,使

$$\sum_{j=1}^{r} d_j\boldsymbol{p}_j = \boldsymbol{0}.$$

令 $\boldsymbol{d}=(d_1,d_2,\cdots,d_r,0,0,\cdots,0)^{\mathrm{T}}\in\mathbb{R}^n$,因为 $\bar{x}_j>0(1\leqslant j\leqslant r)$,所以可选取充分小的 $\varepsilon>0$,使

$$\bar{\boldsymbol{x}} + \varepsilon\boldsymbol{d} \geqslant \boldsymbol{0},\bar{\boldsymbol{x}} - \varepsilon\boldsymbol{d} \geqslant \boldsymbol{0}$$

由于

$$\boldsymbol{A}(\bar{\boldsymbol{x}} \pm \varepsilon\boldsymbol{d}) = \boldsymbol{A}\bar{\boldsymbol{x}} \pm \varepsilon\boldsymbol{A}\boldsymbol{d} = \boldsymbol{b},$$

因此 $\bar{\boldsymbol{x}}+\varepsilon\boldsymbol{d},\bar{\boldsymbol{x}}\in K$,而

$$\bar{x} = \frac{1}{2}(\bar{x} + \varepsilon d) + \frac{1}{2}(\bar{x} - \varepsilon d),$$

即 $\bar{x}$ 可表示为 K 中两个相异点的严格凸组合，故 $\bar{x}$ 不是 K 的极点.

反过来，设 $\bar{x}$ 的非零分量对应的列向量 $p_1, p_2, \cdots, p_r$ 线性无关，如果存在 $x_1, x_2 \in K$，使

$$\bar{x} = \lambda_0 x_1 + (1 - \lambda_0) x_2, 0 < \lambda_0 < 1,$$

则由分量 $\bar{x}_j = 0 (r+1 \leqslant j \leqslant n)$ 可知

$$x_j^{(1)} = x_j^{(2)} = 0, r+1 \leqslant j \leqslant n.$$

又由 $Ax_1 = b$ 和 $Ax_2 = b$ 可知

$$\sum_{j=1}^{r} x_j^{(1)} p_j = b, \sum_{j=1}^{r} x_j^{(2)} p_j = b,$$

即

$$\sum_{j=1}^{r} (x_j^{(1)} - x_j^{(2)}) p_j = 0.$$

再由 $p_1, p_2, \cdots, p_r$ 线性无关，可知

$$x_j^{(1)} = x_j^{(2)}, 1 \leqslant j \leqslant r,$$

从而 $x_1 = x_2$，所以 $\bar{x}$ 为 K 的极点. □

根据这个定理，K 的极点个数不超过矩阵 A 中线性无关的列向量组的个数，所以我们有

推论 2.2.3 K 的极点个数是有限的. □

当 $K \neq \varnothing$ 时，设 K 的全部极点为 $x_1, x_2, \cdots, x_k, k \geqslant 1$，记有限点集 $\{x_1, x_2, \cdots, x_k\}$ 的凸包为 $\hat{K}$，即

$$\hat{K} = \left\{ \sum_{i=1}^{k} \lambda_i x_i \,\middle|\, \lambda_i \geqslant 0 (1 \leqslant i \leqslant k), \sum_{i=1}^{k} \lambda_i = 1 \right\}.$$

又记

$$K_0 = \{ y \in \mathbb{R}^n \mid Ay = 0, y \geqslant 0 \},$$

则显然 $0 \in K_0$，即 $K_0 \neq \varnothing$.

定理 2.2.4 若 $K \neq \varnothing$，则

$$K = \hat{K} + K_0 = \{ x + y \mid x \in \hat{K}, y \in K_0 \}.$$

证明 先证 $K \supseteq \hat{K} + K_0$.

任取 $x \in \hat{K} + K_0$，则存在 $\lambda_i \geqslant 0 (1 \leqslant i \leqslant k)$，$\sum_{i=1}^{k} \lambda_i = 1$ 及 $y \in K_0$，使

$$x = \sum_{i=1}^{k} \lambda_i x_i + y.$$

从而 $x \geqslant 0$，且

$$Ax = \sum_{i=1}^{k} \lambda_i Ax_i + Ay = \sum_{i=1}^{k} \lambda_i b = b,$$

这表明 $x \in K$.

再证 $K \subseteq \hat{K} + K_0$.

设 $x \in K$,对 x 的非零分量个数用归纳法证明 $x \in \hat{K} + K_0$.

当 x 的非零分量个数为 0 时,即 $x = 0$,由极点的定义,$x = 0$ 为 K 的极点,从而

$$x = 0 + 0 \in \hat{K} + K_0.$$

现在假设 x 的非零分量个数小于 r 时,$x \in \hat{K} + K_0$. 下设 x 的非零分量个数等于 $r(1 \leqslant r \leqslant n)$,不失一般性,设 x 的非零分量为 $x_1, x_2, \cdots, x_r$,它们对应的列向量为 $p_1, p_2, \cdots, p_r$. 若 $p_1, p_2, \cdots, p_r$ 线性无关,则由定理 2.2.2 知,x 是 K 的极点,显然 $x = x + 0 \in \hat{K} + K_0$. 若 $p_1, p_2, \cdots, p_r$ 线性相关,则必存在不全为零的数 d_1, $d_2, \cdots, d_r$(不妨设其中至少有一个正数),使

$$\sum_{i=1}^{r} d_i p_i = 0.$$

记 $d = (d_1, d_2, \cdots, d_r, 0, 0, \cdots, 0)^{\mathrm{T}} \in \mathbb{R}^n$,并令

$$\varepsilon_1 = \min\left\{ \frac{x_j}{d_j} \,\middle|\, d_j > 0, 1 \leqslant j \leqslant r \right\},$$

$$\bar{x} = x - \varepsilon_1 d,$$

则 $\varepsilon_1 > 0$,$\bar{x} \in K$,且 $\bar{x}$ 的非零分量个数小于 r,从而由归纳假设知,$\bar{x} \in \hat{K} + K_0$,即存在 $\bar{\lambda}_i \geqslant 0 (1 \leqslant i \leqslant k)$, $\sum_{i=1}^{k} \bar{\lambda}_i = 1$ 及 $\bar{y} \in K_0$,使

$$\bar{x} = \sum_{i=1}^{k} \bar{\lambda}_i x_i + \bar{y}. \tag{2.2.3}$$

下面分两种情形讨论.

(1) 若 $d \geqslant 0$,则由 d 的取法知 $d \neq 0$, $Ad = 0$,即 $d \in K_0$,又由(2.2.3)式知

$$x = \bar{x} + \varepsilon_1 d = \sum_{i=1}^{k} \bar{\lambda}_i x_i + (\bar{y} + \varepsilon_1 d),$$

并注意到 $\varepsilon_1 > 0$, $\bar{y} \in K_0$, $d \in K_0$,因此由 K_0 的定义知 $\bar{y} + \varepsilon_1 d \in K_0$,从而 $x \in \hat{K} + K_0$.

(2) 若 $d \ngeqslant 0$,令

$$\varepsilon_2 = \max\left\{ \frac{x_j}{d_j} \,\middle|\, d_j < 0, 1 \leqslant j \leqslant r \right\},$$

$$\tilde{x} = x - \varepsilon_2 d,$$

则 $\varepsilon_2 < 0$,$\tilde{\boldsymbol{x}} \in K$,且 $\tilde{\boldsymbol{x}}$ 的非零分量个数小于 r,于是由归纳假设知,$\tilde{\boldsymbol{x}} \in \hat{K} + K_0$,即存在 $\tilde{\lambda}_i \geqslant 0 (1 \leqslant i \leqslant k)$,$\sum_{i=1}^{k} \tilde{\lambda}_i = 1$ 及 $\tilde{\boldsymbol{y}} \in K_0$,使

$$\tilde{\boldsymbol{x}} = \sum_{i=1}^{k} \tilde{\lambda}_i \boldsymbol{x}_i + \tilde{\boldsymbol{y}}. \tag{2.2.4}$$

若记 $\alpha = \dfrac{-\varepsilon_2}{\varepsilon_1 - \varepsilon_2}$,则 $0 < \alpha < 1$,且

$$\boldsymbol{x} = \frac{-\varepsilon_2}{\varepsilon_1 - \varepsilon_2} \bar{\boldsymbol{x}} + \frac{\varepsilon_1}{\varepsilon_1 - \varepsilon_2} \tilde{\boldsymbol{x}} = \alpha \bar{\boldsymbol{x}} + (1 - \alpha) \tilde{\boldsymbol{x}}.$$

从而由(2.2.3)和(2.2.4)两式,得

$$\boldsymbol{x} = \sum_{i=1}^{k} (\alpha \bar{\lambda}_i + (1 - \alpha) \tilde{\lambda}_i) \boldsymbol{x}_i + (\alpha\, \bar{\boldsymbol{y}} + (1 - \alpha) \tilde{\boldsymbol{y}}),$$

并且

$$\alpha \bar{\lambda}_i + (1 - \alpha) \tilde{\lambda}_i \geqslant 0, 1 \leqslant i \leqslant k,$$

$$\sum_{i=1}^{k} (\alpha \bar{\lambda}_i + (1 - \alpha) \tilde{\lambda}_i) = \alpha + (1 - \alpha) = 1.$$

又易知 $\alpha\, \bar{\boldsymbol{y}} + (1 - \alpha) \tilde{\boldsymbol{y}} \in K_0$,因此 $\boldsymbol{x} \in \hat{K} + K_0$. □

推论 2.2.5 设 $K \neq \varnothing$,则 K 有界的充要条件是 $K_0 = \{\boldsymbol{0}\}$.

证明 若 $K_0 = \{\boldsymbol{0}\}$,则由定理 2.2.4,对于任意 $\boldsymbol{x} \in K$,存在 $\lambda_i \geqslant 0 (1 \leqslant i \leqslant k)$,$\sum_{i=1}^{k} \lambda_i = 1$,使 $\boldsymbol{x} = \sum_{i=1}^{k} \lambda_i \boldsymbol{x}_i$. 取

$$M_0 = \max_{1 \leqslant i \leqslant k} \| \boldsymbol{x}_i \|^2 + 1,$$

从而

$$\| \boldsymbol{x} \|^2 \leqslant \sum_{i=1}^{k} \lambda_i^2 \| \boldsymbol{x}_i \|^2 \leqslant \sum_{i=1}^{k} \lambda_i M_0 = M_0, \forall \boldsymbol{x} \in K,$$

即 K 有界.

反过来,若存在 $\bar{\boldsymbol{y}} \in K_0$,$\bar{\boldsymbol{y}} \neq \boldsymbol{0}$,即存在 $1 \leqslant t \leqslant n$,使 $\tilde{y}_t > 0$,则对于任意 $M > 0$,取 $\tilde{\boldsymbol{x}} \in \hat{K}$,并令

$$\bar{\boldsymbol{x}} = \tilde{\boldsymbol{x}} + \left(\frac{M}{\tilde{y}_t} + 1 \right) \tilde{\boldsymbol{y}},$$

则易知 $\bar{\boldsymbol{x}} \in K$,且

$$\bar{x}_t = \tilde{x}_t + \left(\frac{M}{\tilde{y}_t} + 1 \right) \tilde{y}_t = \tilde{x}_t + M + \tilde{y}_t > M,$$

即知 K 无界. □

由这个推论立即得到下面的一个重要结论.

推论 2.2.6(多面体表示定理) 若 K 是非空有界集, $\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_k$ 是 K 的全部极点,则 $\boldsymbol{x}\in K$ 的充要条件是存在 $\lambda_i\geqslant 0(1\leqslant i\leqslant k)$, $\sum_{i=1}^{k}\lambda_i=1$,使

$$\boldsymbol{x}=\sum_{i=1}^{k}\lambda_i\boldsymbol{x}_i. \qquad \square$$

这说明多面体就是由其全部极点所生成的凸包.类似于多面体表示定理,下面给出多胞形表示定理,为此引进凸集的方向和极方向的概念.

定义 2.2.2 设 $S\subseteq\mathbb{R}^n$ 是非空凸集, $\boldsymbol{d}\in\mathbb{R}^n$, $\boldsymbol{d}\neq\boldsymbol{0}$,如果

$$\boldsymbol{x}+\lambda\boldsymbol{d}\in S,\ \forall\boldsymbol{x}\in S,\ \forall\lambda>0,$$

则称 $\boldsymbol{d}$ 是 S 的一个方向(direction).又设 $\boldsymbol{d}_1$ 和 $\boldsymbol{d}_2$ 是 S 的两个方向,若存在 $\alpha>0$,使 $\boldsymbol{d}_1=\alpha\boldsymbol{d}_2$,则称 $\boldsymbol{d}_1$ 和 $\boldsymbol{d}_2$ 是相同的方向.如果 S 中的方向 $\boldsymbol{d}$ 不能表示为两个不相同的方向的正的线性组合,则称 $\boldsymbol{d}$ 为 S 的极方向(extreme direction).

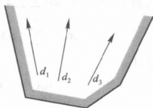


图 2.2.1 极方向与非极方向

在图 2.2.1 中,方向 $\boldsymbol{d}_1$ 和 $\boldsymbol{d}_3$ 是极方向,而方向 $\boldsymbol{d}_2$ 不是极方向.

定理 2.2.7 设 $K\neq\varnothing$,则 $\boldsymbol{d}$ 为 K 的方向当且仅当

$$\boldsymbol{d}\geqslant\boldsymbol{0},\boldsymbol{d}\neq\boldsymbol{0},\boldsymbol{A}\boldsymbol{d}=\boldsymbol{0}. \tag{2.2.5}$$

证明 若 $\boldsymbol{d}$ 满足(2.2.5)式,则对于任意 $\boldsymbol{x}\in K$ 及 $\lambda>0$,有 $\boldsymbol{x}+\lambda\boldsymbol{d}\geqslant\boldsymbol{0}$,且

$$\boldsymbol{A}(\boldsymbol{x}+\lambda\boldsymbol{d})=\boldsymbol{A}\boldsymbol{x}+\lambda\boldsymbol{A}\boldsymbol{d}=\boldsymbol{b},$$

即 $\boldsymbol{x}+\lambda\boldsymbol{d}\in K$,故 $\boldsymbol{d}$ 为 K 的方向.

反过来,若 $\boldsymbol{d}$ 是 K 的方向,则 $\boldsymbol{d}\neq\boldsymbol{0}$,且对任意 $\boldsymbol{x}\in K$ 及 $\lambda>0$,有 $\boldsymbol{x}+\lambda\boldsymbol{d}\in K$,即

$$\boldsymbol{A}(\boldsymbol{x}+\lambda\boldsymbol{d})=\boldsymbol{b},\boldsymbol{x}+\lambda\boldsymbol{d}\geqslant\boldsymbol{0}.$$

因 $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{b}$, $\boldsymbol{x}\geqslant\boldsymbol{0}$, λ 可取任意正数,故

$$\boldsymbol{A}\boldsymbol{d}=\boldsymbol{0},\boldsymbol{d}\geqslant\boldsymbol{0},$$

即(2.2.5)式成立. $\square$

由定理 2.2.7 可知,当 $K\neq\varnothing$ 时, $\boldsymbol{d}$ 为 K 的方向当且仅当 $\boldsymbol{d}\in K_0$ 且 $\boldsymbol{d}\neq\boldsymbol{0}$,于是由推论 2.2.5 即知, K 为有界集当且仅当 K 没有方向.

为了讨论 K 的极方向的特征,我们记

$$\bar{K}_0=\{\boldsymbol{y}\mid\boldsymbol{A}\boldsymbol{y}=\boldsymbol{0},\boldsymbol{e}^{\mathrm{T}}\boldsymbol{y}=1,\boldsymbol{y}\geqslant\boldsymbol{0}\},$$

其中 $\boldsymbol{e}=(1,1,\cdots,1)^{\mathrm{T}}\in\mathbb{R}^n$.

根据定理 2.2.7 和 $\bar{K}_0$ 的定义,易知 $\boldsymbol{d}$ 为 K 的方向当且仅当 $\dfrac{\boldsymbol{d}}{\boldsymbol{e}^{\mathrm{T}}\boldsymbol{d}}\in\bar{K}_0$. 并且还

可以得到

引理 2.2.8 $\boldsymbol{d}$ 为 K 的极方向当且仅当$\frac{\boldsymbol{d}}{\boldsymbol{e}^{\mathrm{T}}\boldsymbol{d}}$为 $\bar{K}_0$ 的极点.

证明 必要性. 设 $\boldsymbol{d}$ 为 K 的极方向,由上知,$\tilde{\boldsymbol{d}}=\frac{\boldsymbol{d}}{\boldsymbol{e}^{\mathrm{T}}\boldsymbol{d}}\in\bar{K}_0$. 如果存在 $\boldsymbol{y}_1,\boldsymbol{y}_2\in\bar{K}_0$,及 $\lambda\in(0,1)$,使得

$$\tilde{\boldsymbol{d}}=\lambda\boldsymbol{y}_1+(1-\lambda)\boldsymbol{y}_2,$$

则

$$\boldsymbol{d}=\boldsymbol{e}^{\mathrm{T}}\boldsymbol{d}(\lambda\boldsymbol{y}_1+(1-\lambda)\boldsymbol{y}_2).$$

由于 $\boldsymbol{d}$ 是 K 的极方向,且 $\lambda\boldsymbol{e}^{\mathrm{T}}\boldsymbol{d}>0,(1-\lambda)\boldsymbol{e}^{\mathrm{T}}\boldsymbol{d}>0$,因此 $\boldsymbol{y}_1$ 与 $\boldsymbol{y}_2$ 必是两个相同的方向,即存在 $\alpha>0$,使得 $\boldsymbol{y}_1=\alpha\boldsymbol{y}_2$. 于是有

$$\boldsymbol{d}=\boldsymbol{e}^{\mathrm{T}}\boldsymbol{d}(\lambda\alpha\boldsymbol{y}_2+(1-\lambda)\boldsymbol{y}_2).$$

由 $\boldsymbol{y}_2\in\bar{K}_0$ 知 $\boldsymbol{e}^{\mathrm{T}}\boldsymbol{y}_2=1$,从而把上式两边同时左乘 $\boldsymbol{e}^{\mathrm{T}}$,得

$$1=(\lambda\alpha+(1-\lambda))\boldsymbol{e}^{\mathrm{T}}\boldsymbol{y}_2=\lambda(\alpha-1)+1,$$

所以 $\alpha=1$,即 $\boldsymbol{y}_1=\boldsymbol{y}_2$,这表明 $\tilde{\boldsymbol{d}}$ 是 $\bar{K}_0$ 的极点.

充分性. 设$\frac{\boldsymbol{d}}{\boldsymbol{e}^{\mathrm{T}}\boldsymbol{d}}$为 $\bar{K}_0$ 的极点,显然 $\boldsymbol{d}$ 为 K 的方向. 假若 $\boldsymbol{d}$ 不是 K 的极方向,则 K 中存在两个不相同的方向 $\boldsymbol{y}_1$ 和 $\boldsymbol{y}_2$,及 $\beta_1,\beta_2>0$,使得

$$\boldsymbol{d}=\beta_1\boldsymbol{y}_1+\beta_2\boldsymbol{y}_2.$$

因 $\boldsymbol{y}_1$ 和 $\boldsymbol{y}_2$ 是 K 的两个不相同的方向,故由定理 2.2.7 知

$$\boldsymbol{y}_1\geqslant\boldsymbol{0},\boldsymbol{y}_2\geqslant\boldsymbol{0},\boldsymbol{y}_1\neq\boldsymbol{0},\boldsymbol{y}_2\neq\boldsymbol{0},\boldsymbol{y}_1\neq\boldsymbol{y}_2,$$

所以 $\boldsymbol{e}^{\mathrm{T}}\boldsymbol{y}_1>0,\boldsymbol{e}^{\mathrm{T}}\boldsymbol{y}_2>0$,并且

$$1=\frac{\boldsymbol{e}^{\mathrm{T}}\boldsymbol{d}}{\boldsymbol{e}^{\mathrm{T}}\boldsymbol{d}}=\frac{\beta_1\boldsymbol{e}^{\mathrm{T}}\boldsymbol{y}_1}{\boldsymbol{e}^{\mathrm{T}}\boldsymbol{d}}+\frac{\beta_2\boldsymbol{e}^{\mathrm{T}}\boldsymbol{y}_2}{\boldsymbol{e}^{\mathrm{T}}\boldsymbol{d}},$$

于是

$$\frac{\boldsymbol{d}}{\boldsymbol{e}^{\mathrm{T}}\boldsymbol{d}}=\frac{\beta_1\boldsymbol{e}^{\mathrm{T}}\boldsymbol{y}_1}{\boldsymbol{e}^{\mathrm{T}}\boldsymbol{d}}\frac{\boldsymbol{y}_1}{\boldsymbol{e}^{\mathrm{T}}\boldsymbol{y}_1}+\frac{\beta_2\boldsymbol{e}^{\mathrm{T}}\boldsymbol{y}_2}{\boldsymbol{e}^{\mathrm{T}}\boldsymbol{d}}\frac{\boldsymbol{y}_2}{\boldsymbol{e}^{\mathrm{T}}\boldsymbol{y}_2},$$

这说明$\frac{\boldsymbol{d}}{\boldsymbol{e}^{\mathrm{T}}\boldsymbol{d}}$可以表示为 $\bar{K}_0$ 中两个不同的点的严格凸组合,此与$\frac{\boldsymbol{d}}{\boldsymbol{e}^{\mathrm{T}}\boldsymbol{d}}$为 $\bar{K}_0$ 的极点相矛盾.

根据推论 2.2.5,K 为无界集当且仅当 $K_0\neq\{\boldsymbol{0}\}$,而后者又等价于 $\bar{K}_0\neq\varnothing$,即 $\bar{K}_0$ 为多面体. 因为 $\bar{K}_0$ 与 K 具有相同的形式,所以由定理 2.2.1 和推论 2.2.3 可知,多面体 $\bar{K}_0$ 存在有限个极点. 再由引理 2.2.8 知,当 K 为无界集时,K 存在有限个极方向. 由此可以给出集合 K_0 的另一种表示.

引理 2.2.9 设 K 为无界集,且 K 的全部极方向为 $\boldsymbol{d}_1,\boldsymbol{d}_2,\cdots,\boldsymbol{d}_l$,则必有

$$K_0 = \left\{ \sum_{j=1}^{l} \mu_j \boldsymbol{d}_j \,\middle|\, \mu_j \geqslant 0, j = 1,2,\cdots,l \right\}.$$

证明 记 $\tilde{K}_0 = \left\{ \sum\limits_{j=1}^{l} \mu_j \boldsymbol{d}_j \,\middle|\, \mu_j \geqslant 0, j=1,2,\cdots,l \right\}$.

因 $\boldsymbol{d}_j$ 为 K 的极方向，故 $\boldsymbol{d}_j \in K_0, j=1,2,\cdots,l$. 从而对一切 $\mu_j \geqslant 0, j=1,2,\cdots,l$，有 $\sum\limits_{j=1}^{l} \mu_j \boldsymbol{d}_j \in K_0$，即 $\tilde{K}_0 \subseteq K_0$.

另一方面，任取 $\boldsymbol{y} \in K_0$，若 $\boldsymbol{y}=\boldsymbol{0}$，显然 $\boldsymbol{y} \in \tilde{K}_0$. 若 $\boldsymbol{y} \neq \boldsymbol{0}$，则令

$$\tilde{\boldsymbol{y}} = \frac{\boldsymbol{y}}{\boldsymbol{e}^{\mathrm{T}}\boldsymbol{y}}, \tilde{\boldsymbol{d}}_j = \frac{\boldsymbol{d}_j}{\boldsymbol{e}^{\mathrm{T}}\boldsymbol{d}_j} (j = 1,2,\cdots,l),$$

则 $\tilde{\boldsymbol{y}} \in \bar{K}_0$. 根据前面的分析，当 K 为无界集时，$\bar{K}_0$ 是多面体，且由引理 2.2.8 知，$\tilde{\boldsymbol{d}}_1, \tilde{\boldsymbol{d}}_2, \cdots, \tilde{\boldsymbol{d}}_l$ 为 $\bar{K}_0$ 的全部极点. 因此由推论 2.2.6 可知，存在 $\lambda_j \geqslant 0 (j=1,2,\cdots,l)$，$\sum\limits_{j=1}^{l} \lambda_j = 1$，使得 $\tilde{\boldsymbol{y}} = \sum\limits_{j=1}^{l} \lambda_j \tilde{\boldsymbol{d}}_j$. 于是

$$\boldsymbol{y} = (\boldsymbol{e}^{\mathrm{T}}\boldsymbol{y})\tilde{\boldsymbol{y}} = \sum_{j=1}^{l} \frac{\lambda_j \boldsymbol{e}^{\mathrm{T}}\boldsymbol{y}}{\boldsymbol{e}^{\mathrm{T}}\boldsymbol{d}_j}\boldsymbol{d}_j, \frac{\lambda_j \boldsymbol{e}^{\mathrm{T}}\boldsymbol{y}}{\boldsymbol{e}^{\mathrm{T}}\boldsymbol{d}_j} \geqslant 0 (j = 1,2,\cdots,l),$$

从而 $\boldsymbol{y} \in \tilde{K}_0$，即 $K_0 \subseteq \tilde{K}_0$. 这就证明了 $K_0 = \tilde{K}_0$. □

引理 2.2.9 表明，当 K 是无界时，K 的任一方向均落在由 K 的极方向所生成的凸锥中.

定理 2.2.10（多胞形表示定理） 若 K 为无界集，K 的全部极点为 $\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_k (k \geqslant 1)$，$K$ 的全部极方向为 $\boldsymbol{d}_1, \boldsymbol{d}_2, \cdots, \boldsymbol{d}_l (l \geqslant 1)$，则 $\boldsymbol{x} \in K$ 的充要条件是存在 $\lambda_i \geqslant 0 (i=1,2,\cdots,k)$，$\sum\limits_{i=1}^{k} \lambda_i = 1$ 和 $\mu_j \geqslant 0 (j=1,2,\cdots,l)$，使

$$\boldsymbol{x} = \sum_{i=1}^{k} \lambda_i \boldsymbol{x}_i + \sum_{j=1}^{l} \mu_j \boldsymbol{d}_j.$$

证明 由定理 2.2.4 和引理 2.2.9 即知. □

最后再介绍一种重要的多面体.

由推论 2.2.6，多面体其实就是它的所有极点所生成的凸包. 因此，我们得到多面体的另一种定义.

定义 2.2.3 设 $\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_k \in \mathbb{R}^n$，则称由 $\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_k$ 所生成的凸包为$\mathbb{R}^n$中的多面体. 如果 $\boldsymbol{x}_2 - \boldsymbol{x}_1, \boldsymbol{x}_3 - \boldsymbol{x}_1, \cdots, \boldsymbol{x}_k - \boldsymbol{x}_1$ 线性无关，则称$\mathbb{R}^n$ 中的多面体 $H(\{\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_k\})$ 是$\mathbb{R}^n$ 中以 $\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_k$ 为顶点（vertex）的单纯形（simplex）.

显然，$\mathbb{R}^n$ 中单纯形的顶点个数不超过 $n+1$.

定理 2.2.11 设 $\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_k \in \mathbb{R}^n$，则

（1）多面体 $H(\{\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_k\})$ 中的极点必定是某个 $\boldsymbol{x}_r (1 \leqslant r \leqslant k)$；

（2）如果 $H(\{\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_k\})$ 是单纯形，那么它的极点的全体就是顶点的

全体.

证明 (1) 设 $\bar{\boldsymbol{x}}\in H(\{\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_k\})$, $\bar{\boldsymbol{x}}\neq\boldsymbol{x}_i(i=1,2,\cdots,k)$, 则存在 $0\leqslant\lambda_i<1(i=1,2,\cdots,k)$, $\sum_{i=1}^{k}\lambda_i=1$, 使 $\bar{\boldsymbol{x}}=\sum_{i=1}^{k}\lambda_i\boldsymbol{x}_i$. 设 $\lambda_t>0(1\leqslant t\leqslant k)$, 记

$$\mu_i=\frac{\lambda_i}{1-\lambda_t}(i\neq t),\tilde{\boldsymbol{x}}=\sum_{i\neq t}\mu_i\boldsymbol{x}_i,$$

则 $\mu_i\geqslant0(i\neq t)$, $\sum_{i\neq t}\mu_i=1$, 从而 $\tilde{\boldsymbol{x}}\in H(\{\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_k\})$, 且

$$\bar{\boldsymbol{x}}=\lambda_t\boldsymbol{x}_t+(1-\lambda_t)\tilde{\boldsymbol{x}}.$$

于是由 $\bar{\boldsymbol{x}}\neq\boldsymbol{x}_t$ 知 $\tilde{\boldsymbol{x}}\neq\boldsymbol{x}_t$, 即 $\bar{\boldsymbol{x}}$ 不是极点, 从而(1)得证.

(2) 由(1)知, 单纯形 $H(\{\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_k\})$ 的极点必是它的顶点. 反之, 任取 $\boldsymbol{x}_t(1\leqslant t\leqslant k)$, 要证 $\boldsymbol{x}_t$ 是单纯形 $H(\{\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_k\})$ 的极点, 只要证明: 若存在 $\alpha\in(0,1)$ 及 $\boldsymbol{y},\boldsymbol{z}\in H(\{\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_k\})$, 使

$$\boldsymbol{x}_t=\alpha\boldsymbol{y}+(1-\alpha)\boldsymbol{z},$$

则 $\boldsymbol{y}=\boldsymbol{z}$.

因为 $\boldsymbol{y},\boldsymbol{z}\in H(\{\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_k\})$, 所以

$$\boldsymbol{y}=\sum_{i=1}^{k}\bar{\lambda}_i\boldsymbol{x}_i,\bar{\lambda}_i\geqslant0(i=1,2,\cdots,k),\sum_{i=1}^{k}\bar{\lambda}_i=1,$$

$$\boldsymbol{z}=\sum_{i=1}^{k}\tilde{\lambda}_i\boldsymbol{x}_i,\tilde{\lambda}_i\geqslant0(i=1,2,\cdots,k),\sum_{i=1}^{k}\tilde{\lambda}_i=1,$$

于是

$$\boldsymbol{x}_t=\sum_{i=1}^{k}(\alpha\bar{\lambda}_i+(1-\alpha)\tilde{\lambda}_i)\boldsymbol{x}_i,\sum_{i=1}^{k}(\alpha\bar{\lambda}_i+(1-\alpha)\tilde{\lambda}_i)=1,$$

从而

$$\sum_{i=1}^{k}(\alpha\bar{\lambda}_i+(1-\alpha)\tilde{\lambda}_i)(\boldsymbol{x}_i-\boldsymbol{x}_1)-(\boldsymbol{x}_t-\boldsymbol{x}_1)=\boldsymbol{0}. \tag{2.2.6}$$

因 $\boldsymbol{x}_2-\boldsymbol{x}_1,\boldsymbol{x}_3-\boldsymbol{x}_1,\cdots,\boldsymbol{x}_k-\boldsymbol{x}_1$ 线性无关, 故上式中 $\boldsymbol{x}_i-\boldsymbol{x}_1(i=2,3,\cdots,k)$ 的系数全为0. 若 $t=1$, 则

$$\alpha\bar{\lambda}_i+(1-\alpha)\tilde{\lambda}_i=0,i=2,3,\cdots,k.$$

即有

$$\alpha\bar{\lambda}_1+(1-\alpha)\tilde{\lambda}_1=1.$$

由于 $0<\alpha<1$, $0\leqslant\bar{\lambda}_1\leqslant1$, $0\leqslant\tilde{\lambda}_1\leqslant1$, 因此 $\bar{\lambda}_1=1$, $\tilde{\lambda}_1=1$, 即知 $\boldsymbol{y}=\boldsymbol{x}_1$, $\boldsymbol{z}=\boldsymbol{x}_1$, 故 $\boldsymbol{y}=\boldsymbol{z}$. 若 $t>1$, 则(2.2.6)式中 $\boldsymbol{x}_t-\boldsymbol{x}_1$ 的系数

$$\alpha\bar{\lambda}_t+(1-\alpha)\tilde{\lambda}_t-1=0,$$

即 $\alpha\bar{\lambda}_t+(1-\alpha)\tilde{\lambda}_t=1$.

同理有 $\bar{\lambda}_t=\tilde{\lambda}_t=1$, 从而 $\boldsymbol{y}=\boldsymbol{z}=\boldsymbol{x}_t$. □

定义 2.2.4 设 $S\subset\mathbb{R}^n$ 为多胞形，$\boldsymbol{x}_1,\boldsymbol{x}_2\in S$ 为两个不同的极点，对于任何 $\boldsymbol{x}_0\in\overline{\boldsymbol{x}_1\boldsymbol{x}_2}$，如果存在 $\boldsymbol{x}_3,\boldsymbol{x}_4\in S$ 使得 $\boldsymbol{x}_0$ 为 $\overline{\boldsymbol{x}_3\boldsymbol{x}_4}$ 的内点，就一定有 $\boldsymbol{x}_3,\boldsymbol{x}_4\in\overline{\boldsymbol{x}_1\boldsymbol{x}_2}$，则称 $\boldsymbol{x}_1$ 和 $\boldsymbol{x}_2$ 是 S 的相邻极点（adjacent extreme point），此时，线段 $\overline{\boldsymbol{x}_1\boldsymbol{x}_2}$ 称为 S 的棱（edge）.

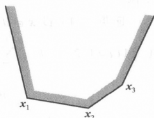


图 2.2.2 多胞形中的相邻极点

在图 2.2.2 所示的多胞形中，$\boldsymbol{x}_1$ 与 $\boldsymbol{x}_2$ 是相邻极点，$\boldsymbol{x}_2$ 与 $\boldsymbol{x}_3$ 也是相邻极点，但极点 $\boldsymbol{x}_1$ 与 $\boldsymbol{x}_3$ 不是相邻的.

定理 2.2.12 单纯形中任何两个相异的顶点都是相邻的极点.

证明 设 $\boldsymbol{x}_p$ 和 $\boldsymbol{x}_q$ 是单纯形 $H(\{\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_k\})$ 的两个相异顶点，则由定理 2.2.11(2) 知 $\boldsymbol{x}_p$ 和 $\boldsymbol{x}_q$ 是两个相异极点. 任取 $\boldsymbol{x}_0\in\overline{\boldsymbol{x}_p\boldsymbol{x}_q}$，即

$$\boldsymbol{x}_0=\alpha\boldsymbol{x}_p+(1-\alpha)\boldsymbol{x}_q,0\leqslant\alpha\leqslant 1, \tag{2.2.7}$$

若存在 $\bar{\boldsymbol{x}},\tilde{\boldsymbol{x}}\in H(\{\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_k\})$，使

$$\boldsymbol{x}_0=\beta\bar{\boldsymbol{x}}+(1-\beta)\tilde{\boldsymbol{x}},0<\beta<1, \tag{2.2.8}$$

则要证 $\bar{\boldsymbol{x}},\tilde{\boldsymbol{x}}\in\overline{\boldsymbol{x}_p\boldsymbol{x}_q}$.

因为 $\bar{\boldsymbol{x}},\tilde{\boldsymbol{x}}\in H(\{\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_k\})$，所以有

$$\bar{\boldsymbol{x}}=\sum_{i=1}^k\bar{\lambda}_i\boldsymbol{x}_i,\bar{\lambda}_i\geqslant 0(i=1,2,\cdots,k),\sum_{i=1}^k\bar{\lambda}_i=1;$$

$$\tilde{\boldsymbol{x}}=\sum_{i=1}^k\tilde{\lambda}_i\boldsymbol{x}_i,\tilde{\lambda}_i\geqslant 0(i=1,2,\cdots,k),\sum_{i=1}^k\tilde{\lambda}_i=1,$$

于是由(2.2.7)和(2.2.8)两式得

$$\beta\sum_{i=1}^k\bar{\lambda}_i\boldsymbol{x}_i+(1-\beta)\sum_{i=1}^k\tilde{\lambda}_i\boldsymbol{x}_i=\alpha\boldsymbol{x}_p+(1-\alpha)\boldsymbol{x}_q,$$

在上式两边减去 $\boldsymbol{x}_1$，整理后得

$$\sum_{i=1}^k(\beta\bar{\lambda}_i+(1-\beta)\tilde{\lambda}_i)(\boldsymbol{x}_i-\boldsymbol{x}_1)-\alpha(\boldsymbol{x}_p-\boldsymbol{x}_1)-(1-\alpha)(\boldsymbol{x}_q-\boldsymbol{x}_1)=\boldsymbol{0}. \tag{2.2.9}$$

由于 $\boldsymbol{x}_2-\boldsymbol{x}_1,\boldsymbol{x}_3-\boldsymbol{x}_1,\cdots,\boldsymbol{x}_k-\boldsymbol{x}_1$ 线性无关，因此上式中 $\boldsymbol{x}_i-\boldsymbol{x}_1(i=2,3,\cdots,k)$ 的系数均为 0，当 $p\neq 1,q\neq 1$ 时，(2.2.9)式中 $\boldsymbol{x}_p-\boldsymbol{x}_1$ 和 $\boldsymbol{x}_q-\boldsymbol{x}_1$ 的系数分别为

$$\beta\bar{\lambda}_p+(1-\beta)\tilde{\lambda}_p-\alpha=0,$$

$$\beta\bar{\lambda}_q+(1-\beta)\tilde{\lambda}_q-(1-\alpha)=0,$$

把上面两式相加，得

$$\beta(\bar{\lambda}_p+\bar{\lambda}_q)+(1-\beta)(\tilde{\lambda}_p+\tilde{\lambda}_q)=1,$$

从而由 $0<\beta<1,0\leqslant\bar{\lambda}_p+\bar{\lambda}_q\leqslant 1$ 和 $0\leqslant\tilde{\lambda}_p+\tilde{\lambda}_q\leqslant 1$ 知

$$\bar{\lambda}_p + \bar{\lambda}_q = 1, \tilde{\lambda}_p + \tilde{\lambda}_q = 1. \tag{2.2.10}$$

当 $p=1$ 时,(2.2.9)式中 $\boldsymbol{x}_i - \boldsymbol{x}_1$ 的系数

$$\beta\bar{\lambda}_i + (1-\beta)\tilde{\lambda}_i = 0, \forall i \neq p, q,$$

从而

$$\beta\bar{\lambda}_p + (1-\beta)\tilde{\lambda}_p + \beta\bar{\lambda}_q + (1-\beta)\tilde{\lambda}_q = 1,$$

即

$$\beta(\bar{\lambda}_p + \bar{\lambda}_q) + (1-\beta)(\tilde{\lambda}_p + \tilde{\lambda}_q) = 1,$$

同样有(2.2.10)式成立. 当 $q=1$ 时,同理可证(2.2.10)式成立.

根据(2.2.10)式,立即得到

$$\bar{\boldsymbol{x}} = \bar{\lambda}_p \boldsymbol{x}_p + \bar{\lambda}_q \boldsymbol{x}_q, \tilde{\boldsymbol{x}} = \tilde{\lambda}_p \boldsymbol{x}_p + \tilde{\lambda}_q \boldsymbol{x}_q,$$

这表明 $\bar{\boldsymbol{x}}, \tilde{\boldsymbol{x}} \in \overline{\boldsymbol{x}_p \boldsymbol{x}_q}$. □

由这个定理我们知道,单纯形 $H(\{\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_k\})$ 恰有 $\binom{k}{2}$ 条棱. 并且 $\mathbb{R}$ 中的单纯形是单点集和线段;$\mathbb{R}^2$ 中的单纯形是单点集、线段和三角形;$\mathbb{R}^3$ 中的单纯形是单点集、线段、三角形和四面体.

$\mathbb{R}^n$ 中具有 $n+1$ 个顶点且所有棱的长都相等的单纯形称为 $\mathbb{R}^n$ 中的正则单纯形(regular simplex). 显然, $\mathbb{R}$ 中正则单纯形是线段, $\mathbb{R}^2$ 中正则单纯形是正三角形, $\mathbb{R}^3$ 中正则单纯形是正四面体.

2.3 凸 函 数

定义 2.3.1 设 $S \subseteq \mathbb{R}^n$ 是非空凸集, $f: S \to \mathbb{R}$, 如果对任意 $\boldsymbol{x}_1, \boldsymbol{x}_2 \in S$ 及 $\lambda \in [0,1]$, 都有

$$f(\lambda \boldsymbol{x}_1 + (1-\lambda)\boldsymbol{x}_2) \leqslant \lambda f(\boldsymbol{x}_1) + (1-\lambda) f(\boldsymbol{x}_2),$$

则称 f 为 S 上的凸函数(convex function). 如果对任意 $\boldsymbol{x}_1, \boldsymbol{x}_2 \in S, \boldsymbol{x}_1 \neq \boldsymbol{x}_2$ 及 $\lambda \in (0,1)$, 都有

$$f(\lambda \boldsymbol{x}_1 + (1-\lambda)\boldsymbol{x}_2) < \lambda f(\boldsymbol{x}_1) + (1-\lambda) f(\boldsymbol{x}_2),$$

则称 f 为 S 上的严格凸函数(strictly convex function).

如果 $-f$ 为 S 上的凸函数,则称 f 为 S 上的凹函数(concave function). 如果 $-f$ 为 S 上的严格凸函数,则称 f 为 S 上的严格凹函数(strictly concave function).

图 2.3.1 给出了凸函数、凹函数和非凸非凹函数的图形. 凸函数的几何解释告诉我们,过凸函数图形上任意两点的弦位于曲线的上方. 而凹函数则恰恰相反.

由凸函数和凹函数的定义易知,线性函数

$$f(\boldsymbol{x}) = \boldsymbol{a}^{\mathrm{T}}\boldsymbol{x} + b, \boldsymbol{a}, \boldsymbol{x} \in \mathbb{R}^n, \boldsymbol{a} \neq \boldsymbol{0}, \ b \in \mathbb{R}$$

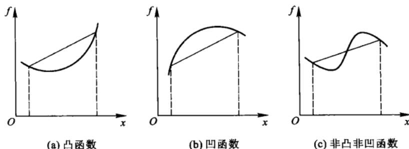


图 2.3.1　凸、凹函数的几何解释

在$\mathbb{R}^n$上既是凸函数也是凹函数.

凸函数具有下列性质.

定理 2.3.1　设f是凸集$S\subseteq\mathbb{R}^n$上的凸函数,$\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_k\in S,\lambda_i\geqslant 0(i=1,2,\cdots,k)$, $\sum_{i=1}^{k}\lambda_i=1$,则

$$f\left(\sum_{i=1}^{k}\lambda_i\boldsymbol{x}_i\right)\leqslant\sum_{i=1}^{k}\lambda_i f(\boldsymbol{x}_i).$$

证明　对k用归纳法即可证明,请读者自行完成. □

定理 2.3.2　设$f_1,f_2,\cdots,f_k$是凸集$S\subseteq\mathbb{R}^n$上的凸函数,则

$$\varphi(\boldsymbol{x})=\sum_{i=1}^{k}\lambda_i f_i(\boldsymbol{x}),\ \forall\lambda_i\geqslant 0(i=1,2,\cdots,k)$$

和

$$\psi(\boldsymbol{x})=\max_{1\leqslant i\leqslant k}f_i(\boldsymbol{x})$$

都是S上的凸函数

证明　对于任意$\boldsymbol{x}_1,\boldsymbol{x}_2\in S$及$\alpha\in(0,1)$,因$f_i$是$S$上的凸函数,故

$$f_i(\alpha\boldsymbol{x}_1+(1-\alpha)\boldsymbol{x}_2)\leqslant\alpha f_i(\boldsymbol{x}_1)+(1-\alpha)f_i(\boldsymbol{x}_2),i=1,2,\cdots,k.$$

从而

$$\begin{aligned}\varphi(\alpha\boldsymbol{x}_1+(1-\alpha)\boldsymbol{x}_2)&=\sum_{i=1}^{k}\lambda_i f_i(\alpha\boldsymbol{x}_1+(1-\alpha)\boldsymbol{x}_2)\\&\leqslant\sum_{i=1}^{k}\lambda_i(\alpha f_i(\boldsymbol{x}_1)+(1-\alpha)f_i(\boldsymbol{x}_2))\\&=\alpha\sum_{i=1}^{k}\lambda_i f_i(\boldsymbol{x}_1)+(1-\alpha)\sum_{i=1}^{k}\lambda_i f_i(\boldsymbol{x}_2)\\&=\alpha\varphi(\boldsymbol{x}_1)+(1-\alpha)\varphi(\boldsymbol{x}_2),\end{aligned}$$

$$\psi(\alpha\boldsymbol{x}_1+(1-\alpha)\boldsymbol{x}_2)=\max_{1\leqslant i\leqslant k}f_i(\alpha\boldsymbol{x}_1+(1-\alpha)\boldsymbol{x}_2)$$

$$\leqslant \max_{1\leqslant i\leqslant k}(\alpha f_i(\boldsymbol{x}_1)+(1-\alpha)f_i(\boldsymbol{x}_2))$$
$$\leqslant \alpha\max_{1\leqslant i\leqslant k}f_i(\boldsymbol{x}_1)+(1-\alpha)\max_{1\leqslant i\leqslant k}f_i(\boldsymbol{x}_2)$$
$$=\alpha\psi(\boldsymbol{x}_1)+(1-\alpha)\psi(\boldsymbol{x}_2),$$

于是 $\varphi(\boldsymbol{x})$ 和 $\psi(\boldsymbol{x})$ 均为 S 上的凸函数. □

为了考察凸集与凸函数的关系,我们引进函数 f 在集合 S 上关于数 α 的水平集(level set)

$$S(f,\alpha)=\{\boldsymbol{x}\in S\,|\,f(\boldsymbol{x})\leqslant\alpha\},$$

其中 $S\subseteq\mathbb{R}^n$, $f:S\to\mathbb{R}$.

定理 2.3.3 设 f 是凸集 $S\subseteq\mathbb{R}^n$ 上的凸函数,则对任意 $\alpha\in\mathbb{R}$,水平集 $S(f,\alpha)$ 是凸集.

证明 对于任意 $\boldsymbol{x}_1,\boldsymbol{x}_2\in S(f,\alpha)$,根据水平集的定义有 $\boldsymbol{x}_1,\boldsymbol{x}_2\in S$,且 $f(\boldsymbol{x}_1)\leqslant\alpha$, $f(\boldsymbol{x}_2)\leqslant\alpha$,由于 S 为凸集,因此,对每个 $\lambda\in(0,1)$,有 $\lambda\boldsymbol{x}_1+(1-\lambda)\boldsymbol{x}_2\in S$. 又因 f 是 S 上的凸函数,故

$$f(\lambda\boldsymbol{x}_1+(1-\lambda)\boldsymbol{x}_2)\leqslant\lambda f(\boldsymbol{x}_1)+(1-\lambda)f(\boldsymbol{x}_2)$$
$$\leqslant\lambda\alpha+(1-\lambda)\alpha=a,$$

从而 $\lambda\boldsymbol{x}_1+(1-\lambda)\boldsymbol{x}_2\in S(f,\alpha)$,故 $S(f,\alpha)$ 是凸集. □

值得注意的是,定理 2.3.3 的逆命题不成立. 例如,当 $S=\{x\in\mathbb{R}\,|\,x\geqslant 0\}$ 时, $f(x)=\sqrt{x}$ 是 S 上的严格凹函数而不是凸函数,但是,对一切 $\alpha\in\mathbb{R}$,水平集 $S(f,\alpha)$ 是凸集.

利用凸函数的定义及有关性质可以判别一个函数是否为凸函数,但有时计算比较复杂,使用很不方便,因此,需要进一步研究凸函数的判别问题.

定理 2.3.4 设 $S\subseteq\mathbb{R}^n$ 是非空开凸集, $f:S\to\mathbb{R}$ 在 S 上可微,则 f 为 S 上的凸函数的充要条件是

$$f(\boldsymbol{x}_2)\geqslant f(\boldsymbol{x}_1)+\nabla f(\boldsymbol{x}_1)^{\mathrm{T}}(\boldsymbol{x}_2-\boldsymbol{x}_1),\ \forall\,\boldsymbol{x}_1,\boldsymbol{x}_2\in S;\tag{2.3.1}$$

f 为 S 上的严格凸函数的充要条件是

$$f(\boldsymbol{x}_2)>f(\boldsymbol{x}_1)+\nabla f(\boldsymbol{x}_1)^{\mathrm{T}}(\boldsymbol{x}_2-\boldsymbol{x}_1),\ \forall\,\boldsymbol{x}_1,\boldsymbol{x}_2\in S,\boldsymbol{x}_1\neq\boldsymbol{x}_2.\tag{2.3.2}$$

证明 必要性. 设 f 是 S 上的凸函数,则对任意 $\boldsymbol{x}_1,\boldsymbol{x}_2\in S$ 及 $\lambda\in(0,1)$,有

$$f(\lambda\boldsymbol{x}_2+(1-\lambda)\boldsymbol{x}_1)\leqslant\lambda f(\boldsymbol{x}_2)+(1-\lambda)f(\boldsymbol{x}_1).$$

于是

$$\frac{f(\boldsymbol{x}_1+\lambda(\boldsymbol{x}_2-\boldsymbol{x}_1))-f(\boldsymbol{x}_1)}{\lambda}\leqslant f(\boldsymbol{x}_2)-f(\boldsymbol{x}_1).$$

因 S 为开集, f 在 S 上可微,故令 $\lambda\to0^+$,得

$$\nabla f(\boldsymbol{x}_1)^{\mathrm{T}}(\boldsymbol{x}_2-\boldsymbol{x}_1)\leqslant f(\boldsymbol{x}_2)-f(\boldsymbol{x}_1),$$

即得(2.3.1)式.

当 f 为 S 上严格凸函数时，对于任意 $\boldsymbol{x}_1,\boldsymbol{x}_2\in S,\boldsymbol{x}_1\neq\boldsymbol{x}_2$，有

$$f\left(\frac{1}{2}\boldsymbol{x}_1+\frac{1}{2}\boldsymbol{x}_2\right)<\frac{1}{2}f(\boldsymbol{x}_1)+\frac{1}{2}f(\boldsymbol{x}_2).$$

记 $\boldsymbol{y}=\frac{1}{2}\boldsymbol{x}_1+\frac{1}{2}\boldsymbol{x}_2$，则由 f 为凸函数的必要条件知

$$f(\boldsymbol{y})\geqslant f(\boldsymbol{x}_1)+\nabla f(\boldsymbol{x}_1)^{\mathrm{T}}(\boldsymbol{y}-\boldsymbol{x}_1),$$

从而

$$\frac{1}{2}f(\boldsymbol{x}_1)+\frac{1}{2}f(\boldsymbol{x}_2)>f(\boldsymbol{x}_1)+\frac{1}{2}\nabla f(\boldsymbol{x}_1)^{\mathrm{T}}(\boldsymbol{x}_2-\boldsymbol{x}_1),$$

即

$$f(\boldsymbol{x}_2)>f(\boldsymbol{x}_1)+\nabla f(\boldsymbol{x}_1)^{\mathrm{T}}(\boldsymbol{x}_2-\boldsymbol{x}_1),$$

故(2.3.2)式成立.

充分性. 设(2.3.1)式成立，则对于任意 $\lambda\in[0,1]$，取 $\boldsymbol{x}=\lambda\boldsymbol{x}_1+(1-\lambda)\boldsymbol{x}_2\in S$，从而

$$f(\boldsymbol{x}_1)\geqslant f(\boldsymbol{x})+\nabla f(\boldsymbol{x})^{\mathrm{T}}(\boldsymbol{x}_1-\boldsymbol{x}),$$
$$f(\boldsymbol{x}_2)\geqslant f(\boldsymbol{x})+\nabla f(\boldsymbol{x})^{\mathrm{T}}(\boldsymbol{x}_2-\boldsymbol{x}),$$

将上述两式分别乘以 λ 和 $(1-\lambda)$ 后，相加得

$$\begin{aligned}\lambda f(\boldsymbol{x}_1)+(1-\lambda)f(\boldsymbol{x}_2)&\geqslant f(\boldsymbol{x})+\nabla f(\boldsymbol{x})^{\mathrm{T}}(\lambda\boldsymbol{x}_1+(1-\lambda)\boldsymbol{x}_2-\boldsymbol{x})\\&=f(\boldsymbol{x})=f(\lambda\boldsymbol{x}_1+(1-\lambda)\boldsymbol{x}_2),\end{aligned}$$

所以 f 为凸函数.

若(2.3.2)式成立，以上各不等式中不等号严格成立，故 f 为严格凸函数. □

这个定理为我们提供了判别一个可微函数是否为凸函数的依据. 它有明显的几何意义，一个可微函数是凸函数的充要条件是函数图形上任一点处的切平面位于曲面的下方.

对于具有二阶连续偏导数的函数，有下面的判别定理.

定理 2.3.5 设 $S\subseteq\mathbb{R}^n$ 是非空开凸集，$f:S\to\mathbb{R}$，在 S 上具有二阶连续偏导数，则 f 是 S 上的凸函数的充要条件是对于一切 $\boldsymbol{x}\in S$，f 在 $\boldsymbol{x}$ 处的 Hesse 矩阵 $\nabla^2 f(\boldsymbol{x})$ 是半正定矩阵.

证明 必要性. 设 f 是 S 上的凸函数，因 S 是开集，故对于任意 $\bar{\boldsymbol{x}}\in S$ 及 $\boldsymbol{x}\in\mathbb{R}^n$，存在 $\delta>0$，使得当 $\lambda\in(0,\delta)$ 时，$\bar{\boldsymbol{x}}+\lambda\boldsymbol{x}\in S$. 由于 f 是 S 上的凸函数，因此由定理 2.3.4，有

$$f(\bar{\boldsymbol{x}}+\lambda\boldsymbol{x})\geqslant f(\bar{\boldsymbol{x}})+\lambda\nabla f(\bar{\boldsymbol{x}})^{\mathrm{T}}\boldsymbol{x}.$$

又因为 f 在 $\bar{\boldsymbol{x}}$ 处具有二阶连续偏导数，所以按 Taylor 公式有

$$f(\bar{\boldsymbol{x}}+\lambda\boldsymbol{x})=f(\bar{\boldsymbol{x}})+\lambda\nabla f(\bar{\boldsymbol{x}})^{\mathrm{T}}\boldsymbol{x}+\frac{1}{2}\lambda^2\boldsymbol{x}^{\mathrm{T}}\nabla^2 f(\bar{\boldsymbol{x}})\boldsymbol{x}+o(\|\lambda\boldsymbol{x}\|^2),$$

从而

$$\frac{1}{2}\lambda^2 \boldsymbol{x}^{\mathrm{T}} \nabla^2 f(\bar{\boldsymbol{x}})\boldsymbol{x} + o(\|\lambda \boldsymbol{x}\|^2) \geqslant 0,$$

将上式两边除以 λ^2 后再令 $\lambda \to 0^+$,得

$$\boldsymbol{x}^{\mathrm{T}} \nabla^2 f(\bar{\boldsymbol{x}})\boldsymbol{x} \geqslant 0,$$

即$\nabla^2 f(\boldsymbol{x})$是半正定的.

充分性. 设$\nabla^2 f(\boldsymbol{x})$在每一点 $\boldsymbol{x} \in S$ 处半正定,由 $f(\boldsymbol{x})$ 在 $\bar{\boldsymbol{x}} \in S$ 处的 Taylor 公式有

$$f(\boldsymbol{x}) = f(\bar{\boldsymbol{x}}) + \nabla f(\bar{\boldsymbol{x}})^{\mathrm{T}}(\boldsymbol{x} - \bar{\boldsymbol{x}}) + \frac{1}{2}(\boldsymbol{x} - \bar{\boldsymbol{x}})^{\mathrm{T}} \nabla^2 f(\hat{\boldsymbol{x}})(\boldsymbol{x} - \bar{\boldsymbol{x}}), \tag{2.3.3}$$

其中 $\hat{\boldsymbol{x}} = \bar{\boldsymbol{x}} + \lambda(\boldsymbol{x} - \bar{\boldsymbol{x}})$, $\lambda \in (0,1)$. 因 S 是凸集,故 $\hat{\boldsymbol{x}} \in S$. 由于$\nabla^2 f(\hat{\boldsymbol{x}})$半正定,因此

$$(\boldsymbol{x} - \bar{\boldsymbol{x}})^{\mathrm{T}} \nabla^2 f(\hat{\boldsymbol{x}})(\boldsymbol{x} - \bar{\boldsymbol{x}}) \geqslant 0,$$

于是由(2.3.3)式有

$$f(\boldsymbol{x}) \geqslant f(\bar{\boldsymbol{x}}) + \nabla f(\bar{\boldsymbol{x}})(\boldsymbol{x} - \bar{\boldsymbol{x}}) \tag{2.3.4}$$

由定理 2.3.4,f 为 S 上的凸函数. □

定理 2.3.6 设 $S \subseteq \mathbb{R}^n$ 是非空开凸集,$f: S \to \mathbb{R}$ 在 S 上具有二阶连续偏导数,如果对每个 $\boldsymbol{x} \in S$,$\nabla^2 f(\boldsymbol{x})$是正定矩阵,则 f 是 S 上的严格凸函数.

证明 因为$\nabla^2 f(\boldsymbol{x})$是正定的,所以当 $\boldsymbol{x} \neq \bar{\boldsymbol{x}}$ 时,(2.3.3)式的末项为正. 从而(2.3.4)式为严格不等式,因此 f 是 S 上的严格凸函数. □

定理 2.3.6 的逆命题不成立. 例如,$f(x) = x^4$ 是 $\mathbb{R}$ 上的严格凸函数,但它的 Hesse 矩阵$\nabla^2 f(x) = 12x^2$ 在点 $x = 0$ 处不是正定的.

定理 2.3.7 设 $f: \mathbb{R}^n \to \mathbb{R}$ 为二次函数,即

$$f(\boldsymbol{x}) = \frac{1}{2}\boldsymbol{x}^{\mathrm{T}}\boldsymbol{Q}\boldsymbol{x} + \boldsymbol{b}^{\mathrm{T}}\boldsymbol{x} + c,$$

其中 $\boldsymbol{Q}$ 是 n 阶对称矩阵,则

(1) f 是$\mathbb{R}^n$ 上的凸函数的充要条件是 $\boldsymbol{Q}$ 为半正定矩阵.

(2) f 是$\mathbb{R}^n$ 上的严格凸函数的充要条件是 $\boldsymbol{Q}$ 为正定矩阵.

证明 由例 1.2.2 知,二次函数 f 在$\mathbb{R}^n$ 上具有二阶连续偏导数,且

$$\nabla f(\boldsymbol{x}) = \boldsymbol{Q}\boldsymbol{x} + \boldsymbol{b}, \nabla^2 f(\boldsymbol{x}) = \boldsymbol{Q},$$

从而由定理 2.3.5,(1)显然成立.

又由定理 2.3.4,f 是$\mathbb{R}^n$ 上的严格凸函数当且仅当

$$f(\boldsymbol{y}) > f(\boldsymbol{x}) + \nabla f(\boldsymbol{x})^{\mathrm{T}}(\boldsymbol{y} - \boldsymbol{x}), \forall \boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n, \boldsymbol{x} \neq \boldsymbol{y}.$$

这等价于

$$f(\boldsymbol{y}) > f(\boldsymbol{x}) + (\boldsymbol{Q}\boldsymbol{x} + \boldsymbol{b})^{\mathrm{T}}(\boldsymbol{y} - \boldsymbol{x}), \forall \boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n, \boldsymbol{x} \neq \boldsymbol{y}.$$

注意到 f 为二次函数,$\boldsymbol{Q}$ 为对称矩阵,因此上式等价于

$$\frac{1}{2}y^{\mathrm{T}}Qy > -\frac{1}{2}x^{\mathrm{T}}Qx + x^{\mathrm{T}}Qy, \forall x,y \in \mathbb{R}^n, x \neq y,$$

这又等价于

$$\frac{1}{2}(y-x)^{\mathrm{T}}Q(y-x) > 0, \forall x,y \in \mathbb{R}^n, x \neq y.$$

这等价于 Q 是正定矩阵. 从而(2)得证. □

2.4 凸 规 划

定义 2.4.1 设 $S \subseteq \mathbb{R}^n$ 为凸集,f 是 S 上的凸函数,则称规划问题

$$\min_{x \in S} f(x) \tag{2.4.1}$$

为凸规划(convex programming)问题.

例 2.4.1 当 f 是 $\mathbb{R}^n$ 上的凸函数时,无约束最优化问题

$$\min_{x \in \mathbb{R}^n} f(x)$$

是凸规划问题. □

例 2.4.2 设 $A \in \mathbb{R}^{m \times n}, b \in \mathbb{R}^m, c \in \mathbb{R}^n$,则线性规划问题

$$\begin{cases} \min & c^{\mathrm{T}}x; \\ \text{s.t.} & Ax = b; \\ & x \geqslant 0 \end{cases} \tag{LP}$$

是凸规划问题. □

证明 因为(LP)的目标函数 $c^{\mathrm{T}}x$ 是线性函数,所以是凸函数. 又由于(LP)的可行域

$$K = \{x \in \mathbb{R}^n \mid Ax = b, x \geqslant 0\}$$

是多胞形,因此 K 是凸集,从而(LP)是凸规划. □

例 2.4.3 设 $S \subseteq \mathbb{R}^n$ 为开凸集,f 是 S 上的凸函数,$g_i(i=1,2,\cdots,m)$ 是 S 上的凹函数,$h_j(j=1,2,\cdots,l)$ 是 $\mathbb{R}^n$ 上的线性函数,则下面三个规划问题

$$\begin{cases} \min & f(x); \\ \text{s.t.} & h_j(x) = 0, j = 1,2,\cdots,l, \end{cases} \tag{2.4.2}$$

$$\begin{cases} \min & f(x); \\ \text{s.t.} & g_i(x) \geqslant 0, i = 1,2,\cdots,m, \end{cases} \tag{2.4.3}$$

$$\begin{cases} \min & f(x); \\ \text{s.t.} & g_i(x) \geqslant 0, i = 1,2,\cdots,m, \\ & h_j(x) = 0, j = 1,2,\cdots,l \end{cases} \tag{2.4.4}$$

都是凸规划问题.

证明 记

$$S_1=\{x\in\mathbb{R}^n \mid h_j(x)=0, j=1,2,\cdots,l\},$$
$$S_2=\{x\in\mathbb{R}^n \mid g_i(x)\geqslant 0, i=1,2,\cdots,m\},$$
$$S_3=S_1\cap S_2.$$

要证问题(2.4.2)、(2.4.3)和(2.4.4)为凸规划问题，只需证明 S_1, S_2 和 S_3 为$\mathbb{R}^n$中的凸集. 因为 $h_j(x)=0(j=1,2,\cdots,l)$ 都是线性函数，所以 S_1 是 l 个超平面的交集，从而 S_1 为凸集. 又由于 $-g_i(x)(i=1,2,\cdots,m)$ 均为凸函数，因此各水平集.

$$S(-g_i,0)=\{x\in\mathbb{R}^n \mid g_i(x)\geqslant 0\}(i=1,2,\cdots,m)$$

都是凸集，于是

$$S_2=\bigcap_{i=1}^{m} S(-g_i,0)$$

也是凸集，从而 $S_3=S_1\cap S_2$ 是凸集. □

凸规划比一般非线性规划重要，是因为它具有下面的基本性质.

定理 2.4.1 凸规划问题(2.4.1)的任何局部极小点都是全局极小点，且它的极小点的集合为凸集.

证明 用反证法证明定理的前一部分. 设 $\bar{x}\in S$ 为问题(2.4.1)的局部极小点，即存在 $\bar{x}$ 的某个 δ 邻域 $N_\delta(\bar{x})$，使

$$f(\bar{x})\leqslant f(x), \forall x\in N_\delta(\bar{x})\cap S.$$

若 $\bar{x}$ 不是问题(2.4.1)的全局极小点，则存在 $\tilde{x}\in S$，使 $f(\tilde{x})<f(\bar{x})$. 由于 f 为 S 上的凸函数，因此对每个 $\lambda\in(0,1)$，有

$$f(\lambda\bar{x}+(1-\lambda)\tilde{x})\leqslant\lambda f(\bar{x})+(1-\lambda)f(\tilde{x})<f(\bar{x}). \qquad (2.4.5)$$

当 λ 充分接近 1 时，可使

$$\lambda\bar{x}+(1-\lambda)\tilde{x}\in N_\delta(\bar{x})\cap S,$$

于是

$$f(\bar{x})\leqslant f(\lambda\bar{x}+(1-\lambda)\tilde{x}),$$

此与(2.4.5)式矛盾. 从而 $\bar{x}$ 是问题(2.4.1)的全局极小点.

由以上证明可知，f 在 S 上的极小值也是它在 S 上的最小值. 设最小值为 α，则问题(2.4.1)的极小点的集合是水平集 $S(f,\alpha)$，由定理 2.3.3，$S(f,\alpha)$ 是凸集. □

定理 2.4.2 在凸规划问题(2.4.1)中，若 f 为 S 上的严格凸函数，且 $\bar{x}$ 为问题(2.4.1)的局部极小点，则 $\bar{x}$ 是问题(2.4.1)的惟一全局极小点.

证明 由定理 2.4.1，$\bar{x}$ 是问题(2.4.1)的全局极小点. 假设 $\tilde{x}$ 也是问题(2.4.1)的全局极小点，且 $\tilde{x}\neq\bar{x}$，则 $f(\tilde{x})=f(\bar{x})$，从而由 f 为 S 上的严格凸函数知

$$f\left(\frac{1}{2}\bar{x}+\frac{1}{2}\tilde{x}\right)<\frac{1}{2}f(\bar{x})+\frac{1}{2}f(\tilde{x})=f(\bar{x}),$$

此与 $\bar{x}$ 为全局极小点相矛盾. 所以问题(2.4.1)的全局极小点必惟一. □

习 题 二

1. 设 $\boldsymbol{A}\in\mathbb{R}^{m\times n}$, $\boldsymbol{b}\in\mathbb{R}^m$, 用定义验证

$$S=\{\boldsymbol{x}\in\mathbb{R}^n \mid \boldsymbol{A}\boldsymbol{x}=\boldsymbol{b},\boldsymbol{x}\geqslant\boldsymbol{0}\}$$

是凸集.

2. 证明定理 2.1.2.

3. 写出下列集合的凸包:

(1) $S_1=\{(0,0)^{\mathrm{T}},(1,0)^{\mathrm{T}},(0,1)^{\mathrm{T}}\}$;

(2) $S_2=\{(x_1,x_2)^{\mathrm{T}} \mid x_1^2\leqslant x_2, x_1^2+x_2^2\leqslant 2\}$.

4. 设 $\boldsymbol{x}_i\in\mathbb{R}^n(i=1,2,\cdots,k)$, $\boldsymbol{d}_j\in\mathbb{R}^n(j=1,2,\cdots,l)$, 证明集合

$$\left\{\sum_{i=1}^{k}\lambda_i\boldsymbol{x}_i+\sum_{j=1}^{l}\mu_j\boldsymbol{d}_j \;\middle|\; \sum_{i=1}^{k}\lambda_i=1,\lambda_i\geqslant 0(i=1,2,\cdots,k),\mu_j\geqslant 0(j=1,2,\cdots,l)\right\}$$

是凸集.

5. 设 $S\subseteq\mathbb{R}^n$ 为凸集, $\boldsymbol{A}\in\mathbb{R}^{m\times n}$, 证明 $R=\{\boldsymbol{A}\boldsymbol{x} \mid \boldsymbol{x}\in S\}$ 为凸集.

6. 设 $\boldsymbol{A}\in\mathbb{R}^{m\times n}$, $\boldsymbol{B}\in\mathbb{R}^{l\times n}$, $\boldsymbol{b}\in\mathbb{R}^n$, 证明下列两个关系式组有且仅有一组有解:

$$(\text{Ⅰ})\ \boldsymbol{A}\boldsymbol{x}\leqslant\boldsymbol{0},\boldsymbol{B}\boldsymbol{x}=\boldsymbol{0},\boldsymbol{b}^{\mathrm{T}}\boldsymbol{x}>0;$$

$$(\text{Ⅱ})\ \boldsymbol{A}^{\mathrm{T}}\boldsymbol{y}+\boldsymbol{B}^{\mathrm{T}}\boldsymbol{z}=\boldsymbol{b},\boldsymbol{y}\geqslant\boldsymbol{0}.$$

7. 设 $\boldsymbol{A}\in\mathbb{R}^{m\times n}$, $\boldsymbol{b}\in\mathbb{R}^n$, 证明下列两个关系式组有且仅有一组有解:

$$(\text{Ⅰ})\ \boldsymbol{A}\boldsymbol{x}\leqslant\boldsymbol{0},\boldsymbol{x}\geqslant\boldsymbol{0},\boldsymbol{b}^{\mathrm{T}}\boldsymbol{x}>0;$$

$$(\text{Ⅱ})\ \boldsymbol{A}^{\mathrm{T}}\boldsymbol{y}\geqslant\boldsymbol{b},\boldsymbol{y}\geqslant\boldsymbol{0}.$$

8. 证明下面不等式组无解:

$$\begin{cases}x_1+3x_2<0,\\3x_1-x_2<0,\\17x_1+11x_2>0.\end{cases}$$

9. 设 $S\subseteq\mathbb{R}^n$ 为凸集, $\boldsymbol{x}_0\in S$, 且 $S\setminus\{\boldsymbol{x}_0\}$ 为凸集, 证明 $\boldsymbol{x}_0$ 是 S 的一个极点.

10. 证明: 超平面 $H=\{\boldsymbol{x}\in\mathbb{R}^n \mid \boldsymbol{a}^{\mathrm{T}}\boldsymbol{x}=\alpha\}$ 没有极点.

11. 设 $\boldsymbol{A}\in\mathbb{R}^{m\times n}$, $S=\{\boldsymbol{x}\in\mathbb{R}^n \mid \boldsymbol{A}\boldsymbol{x}\leqslant\boldsymbol{0}\}$.

(1) 证明 S 是一个凸集且最多有一个极点;

(2) 给出并证明 S 存在一个极点的充要条件, 求出极点.

12. 证明: 任何一个单纯形的顶点集是惟一确定的.

13. 证明: $f(\boldsymbol{x})$ 为$\mathbb{R}^n$ 上的凸函数的充要条件是对于任意给定的 $\boldsymbol{x}_1,\boldsymbol{x}_2\in\mathbb{R}^n$, 函数

$$\varphi(\lambda)=f(\lambda\boldsymbol{x}_1+(1-\lambda)\boldsymbol{x}_2),\ \forall\lambda\in[0,1]$$

是凸函数.

14. 设 $S\subseteq\mathbb{R}^n$ 为非空凸集, $f:S\to\mathbb{R}$, 证明: f 是凸函数的充要条件是集合

$$\left\{\begin{bmatrix} x \\ \alpha \end{bmatrix} \middle| f(x) \leq \alpha, x \in S, \alpha \in \mathbb{R}\right\}$$

为$\mathbb{R}^{n+1}$中的凸集.

15. 判别下列函数的凸性,并说明理由.

(1) $f(x_1,x_2) = x_1^2 + 2x_1x_2 - 10x_1 + 5x_2$;

(2) $f(x_1,x_2) = x_1 e^{-(x_1+x_2)}$;

(3) $f(x_1,x_2,x_3) = x_1x_2 + 2x_1^2 + x_2^2 + 2x_3^2 - 6x_1x_3$;

(4) $f(x_1,x_2,x_3) = x_1^2 + 3x_2^2 + 9x_3^2 - 2x_1x_2 + 6x_2x_3 + 2x_3x_1$.

16. 设 $g_i:\mathbb{R}^n\to\mathbb{R}$ 为凹函数($i=1,2,\cdots,m$),试判定函数

$$f(x) = \sum_{i=1}^{m} [\min\{0, g_i(x)\}]^2$$

的凸性.

17. 判断下列非线性规划问题是否为凸规划问题:

(1) $\begin{cases} \min & 2x_1^2 + x_2^2 - 2x_1x_2; \\ \text{s. t.} & x_1^2 + x_2^2 - 4 \leqslant 0, \\ & x_1 + x_2 - 2 \leqslant 0, \\ & x_1 - x_2 - 1 \leqslant 0. \end{cases}$

(2) $\begin{cases} \min & 100x_1 + \dfrac{200}{x_1x_2}; \\ \text{s. t.} & -x_1^2 + x_2 \geqslant 0, \\ & x_1 + x_2 \leqslant 4, \\ & x_1 \geqslant 0, x_2 \geqslant 0. \end{cases}$

(3) $\begin{cases} \min & (x_1-3)^2 + (x_2-2)^2; \\ \text{s. t.} & x_1^2 + x_2 = 5, \\ & x_1 + 2x_2 \leqslant 4. \end{cases}$

18. 设 $S\subseteq\mathbb{R}^n$ 为非空凸集, $f:S\to\mathbb{R}$ 是具有一阶连续偏导数的凸函数,证明:$\bar{x}$ 是问题 $\min\limits_{x\in S} f(x)$的最优解的充要条件是

$$\nabla f(\bar{x})^{T}(x - \bar{x}) \geqslant 0, \forall x \in S.$$

第三章 最优性条件

所谓最优性条件,是指最优化问题的最优解所要满足的必要条件或充分条件.这些条件对于最优化算法的建立和最优化理论的推证都是至关重要的.

本章先介绍无约束最优化问题的最优性条件,然后着重介绍约束最优化问题的最优性条件.

3.1 无约束最优化问题的最优性条件

考虑无约束最优化问题

$$\min f(\boldsymbol{x}), \tag{3.1.1}$$

其中 $f:\mathbb{R}^n\to\mathbb{R}$. 这是一个古典的极值问题,在微积分学中已经有所研究,现在对它进一步讨论.

首先,我们给出问题(3.1.1)的局部极小点的一阶必要条件.

定理 3.1.1 设 $f:\mathbb{R}^n\to\mathbb{R}$ 在点 $\bar{\boldsymbol{x}}$ 处可微,若 $\bar{\boldsymbol{x}}$ 是问题(3.1.1)的局部极小点,则

$$\nabla f(\bar{\boldsymbol{x}}) = \boldsymbol{0}.$$

证明 用反证法. 假设 $\nabla f(\bar{\boldsymbol{x}})\neq\boldsymbol{0}$,取 $\boldsymbol{d}=-\nabla f(\bar{\boldsymbol{x}})$,则有

$$\nabla f(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{d} = -\|\nabla f(\bar{\boldsymbol{x}})\|^2 < 0.$$

从而根据定理 1.2.3,$\boldsymbol{d}$ 是 f 在 $\bar{\boldsymbol{x}}$ 处的下降方向,即存在 $\delta>0$,使

$$f(\bar{\boldsymbol{x}}+\lambda\boldsymbol{d}) < f(\bar{\boldsymbol{x}}),\ \forall\,\lambda\in(0,\delta),$$

此与 $\bar{\boldsymbol{x}}$ 为局部极小点矛盾. □

$\mathbb{R}^n$ 中集合 S 的内部(interior)是指

$$\mathrm{int}S = \{\boldsymbol{x}\in\mathbb{R}^n \mid 存在\ N_\delta(\boldsymbol{x})\subseteq S\}.$$

定义 3.1.1 设 $f:S\subseteq\mathbb{R}^n\to\mathbb{R}$ 在 $\bar{\boldsymbol{x}}\in\mathrm{int}\,S$ 处可微,若

$$\nabla f(\bar{\boldsymbol{x}}) = \boldsymbol{0},$$

则称 $\bar{\boldsymbol{x}}$ 为 f 的平稳点(stationary point).

由定理 3.1.1 可知,无约束最优化问题的局部极小点一定是目标函数的平稳点. 反之不然,这是因为,函数的平稳点可以是它的极小点,也可以是极大点,也可以二者都不是(即鞍点)(见图 3.1.1).

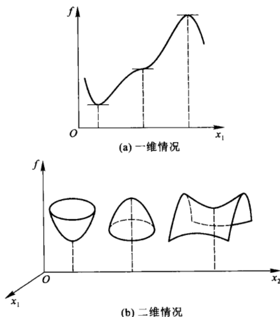


图 3.1.1 平稳点与极值点

下面利用 Hesse 矩阵给出局部极小点的二阶必要条件.

定理 3.1.2 设 $f:\mathbb{R}^n\to\mathbb{R}$ 在点 $\bar{x}\in\mathbb{R}^n$ 处具有二阶连续偏导数. 若 $\bar{x}$ 是问题 (3.1.1) 的局部极小点,则 $\nabla f(\bar{x})=\mathbf{0}$,且 $\nabla^2 f(\boldsymbol{x})$ 半正定.

证明 由定理 3.1.1,$\nabla f(\bar{x})=\mathbf{0}$,故只需证明 Hesse 矩阵 $\nabla^2 f(\bar{x})$ 半正定.

因为 f 在点 $\bar{x}$ 处具有二阶连续偏导数,且 $\nabla f(\bar{x})=\mathbf{0}$,所以由二阶 Taylor 公式,对于任意非零向量 $\boldsymbol{d}\in\mathbb{R}^n$ 和充分小的 $\lambda>0$,有

$$f(\bar{x}+\lambda\boldsymbol{d})=f(\bar{x})+\frac{1}{2}\lambda^2\boldsymbol{d}^{\mathrm{T}}\nabla^2 f(\bar{x})\boldsymbol{d}+o(\|\lambda\boldsymbol{d}\|^2). \tag{3.1.2}$$

由于 $\bar{x}$ 是局部极小点,因此当 λ 充分小时,有

$$f(\bar{x}+\lambda\boldsymbol{d})\geqslant f(\bar{x}).$$

从而由(3.1.2)式知

$$\boldsymbol{d}^{\mathrm{T}}\nabla^2 f(\bar{x})\boldsymbol{d}+2\frac{o(\|\lambda\boldsymbol{d}\|^2)}{\|\lambda\boldsymbol{d}\|^2}\|\boldsymbol{d}\|^2\geqslant 0,$$

令 $\lambda\to 0^+$,得

$$\boldsymbol{d}^{\mathrm{T}}\nabla^2 f(\bar{x})\boldsymbol{d}\geqslant 0,$$

因而 $\nabla^2 f(\bar{x})$ 是半正定的. □

现在给出局部极小点的二阶充分条件.

定理 3.1.3 设$f:\mathbb{R}^n\to\mathbb{R}$在点$\bar{x}\in\mathbb{R}^n$处具有二阶连续偏导数. 若$\nabla f(\bar{x})=\mathbf{0}$,且$\nabla^2 f(\bar{x})$正定,则$\bar{x}$是问题(3.1.1)的严格局部极小点.

证明 因为f在点$\bar{x}$处具有二阶连续偏导数,且$\nabla f(\bar{x})=\mathbf{0}$,所以由二阶Taylor公式,对于任意非零向量$\boldsymbol{d}\in\mathbb{R}^n$和充分小的$\lambda>0$,有

$$f(\bar{x}+\lambda\boldsymbol{d})=f(\bar{x})+\frac{1}{2}\lambda^2\boldsymbol{d}^{\mathrm{T}}\nabla^2 f(\bar{x})\boldsymbol{d}+o(\|\lambda\boldsymbol{d}\|^2).$$

因$\nabla^2 f(\bar{x})$正定,$\boldsymbol{d}\neq\mathbf{0}$,故

$$\boldsymbol{d}^{\mathrm{T}}\nabla^2 f(\bar{x})\boldsymbol{d}>0,$$

所以当λ充分小时,有

$$\frac{1}{2}\lambda^2\boldsymbol{d}^{\mathrm{T}}\nabla^2 f(\bar{x})\boldsymbol{d}+o(\|\lambda\boldsymbol{d}\|^2)>0,$$

即

$$f(\bar{x}+\lambda\boldsymbol{d})>f(\bar{x}).$$

由$\boldsymbol{d}$的任意性知,$\bar{x}$为(3.3.1)严格局部极小点. □

需要指出的是,定理3.1.2不是充分条件,定理3.1.3不是必要条件. 反例分别见例3.1.1和例3.1.2.

例 3.1.1 对于无约束最优化问题

$$\min f(\boldsymbol{x})=x_1^2-x_2^3,$$

其中$\boldsymbol{x}=(x_1,x_2)^{\mathrm{T}}\in\mathbb{R}^2$,显然

$$\nabla f(\boldsymbol{x})=(2x_1,-3x_2^2)^{\mathrm{T}},\ \forall\boldsymbol{x}\in\mathbb{R}^2.$$

令$\nabla f(\boldsymbol{x})=\mathbf{0}$,得$f(\boldsymbol{x})$的平稳点$\bar{x}=(0,0)^{\mathrm{T}}$,而且

$$\nabla^2 f(\boldsymbol{x})=\begin{bmatrix}2 & 0\\ 0 & -6x_2\end{bmatrix},\nabla^2 f(\bar{x})=\begin{bmatrix}2 & 0\\ 0 & 0\end{bmatrix},$$

即$\nabla^2 f(\bar{x})$为半正定矩阵.

但是,在$\bar{x}$的任意δ邻域$\|\boldsymbol{x}-\bar{x}\|<\delta$内,总可以取到$\tilde{x}=\left(0,\frac{\delta}{2}\right)^{\mathrm{T}}$,使$f(\tilde{x})<f(\bar{x})$,即$\bar{x}$不是局部极小点. □

例 3.1.2 对于无约束最优化问题

$$\min f(\boldsymbol{x})=x_1^4+2x_1^2x_2^2+x_2^4,$$

其中$\boldsymbol{x}=(x_1,x_2)^{\mathrm{T}}\in\mathbb{R}^2$,易知

$$\nabla f(\boldsymbol{x})=(4x_1^3+4x_1x_2^2,4x_1^2x_2+4x_2^3)^{\mathrm{T}},$$

从而得平稳点$\bar{x}=(0,0)^{\mathrm{T}}$,并且

$$\nabla^2 f(\boldsymbol{x})=\begin{bmatrix}12x_1^2+4x_2^2 & 8x_1x_2\\ 8x_1x_2 & 4x_1^2+12x_2^2\end{bmatrix},\nabla^2 f(\bar{x})=\begin{bmatrix}0 & 0\\ 0 & 0\end{bmatrix}.$$

即$\nabla^2 f(\bar{x})$不是正定矩阵. 但是$f(\boldsymbol{x})=(x_1^2+x_2^2)^2$在$\bar{x}$处取最小值,即$\bar{x}$为严格局

部极小点. □

例 3.1.3 求解下面无约束最优化问题

$$\min f(\boldsymbol{x}) = \frac{1}{3}x_1^3 + \frac{1}{3}x_2^3 - x_2^2 - x_1,$$

其中 $\boldsymbol{x} = (x_1, x_2)^{\mathrm{T}}$.

解 因为

$$\nabla f(\boldsymbol{x}) = \begin{bmatrix} x_1^2 - 1 \\ x_2^2 - 2x_2 \end{bmatrix}, \nabla^2 f(\boldsymbol{x}) = \begin{bmatrix} 2x_1 & 0 \\ 0 & 2x_2 - 2 \end{bmatrix},$$

所以令 $\nabla f(\boldsymbol{x}) = \boldsymbol{0}$,有

$$\begin{cases} x_1^2 - 1 = 0, \\ x_2^2 - 2x_2 = 0. \end{cases}$$

解此方程组,得到 $f(\boldsymbol{x})$ 的平稳点

$$\boldsymbol{x}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \boldsymbol{x}_2 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \boldsymbol{x}_3 = \begin{bmatrix} -1 \\ 0 \end{bmatrix}, \boldsymbol{x}_4 = \begin{bmatrix} -1 \\ 2 \end{bmatrix}.$$

从而

$$\nabla^2 f(\boldsymbol{x}_1) = \begin{bmatrix} 2 & 0 \\ 0 & -2 \end{bmatrix}, \nabla^2 f(\boldsymbol{x}_2) = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix},$$

$$\nabla^2 f(\boldsymbol{x}_3) = \begin{bmatrix} -2 & 0 \\ 0 & -2 \end{bmatrix}, \nabla^2 f(\boldsymbol{x}_4) = \begin{bmatrix} -2 & 0 \\ 0 & 2 \end{bmatrix}.$$

由于 $\nabla^2 f(\boldsymbol{x}_1)$ 和 $\nabla^2 f(\boldsymbol{x}_4)$ 是不定的,因此 $\boldsymbol{x}_1$ 和 $\boldsymbol{x}_4$ 不是极值点. $\nabla^2 f(\boldsymbol{x}_3)$ 是负定的,故 $\boldsymbol{x}_3$ 不是极小点,实际上它是极大点. $\nabla^2 f(\boldsymbol{x}_2)$ 是正定的,从而 $\boldsymbol{x}_2$ 是严格局部极小点. □

最后,在函数凸性的假设下,给出全局极小点的一阶充分条件.

定理 3.1.4 设 $f:\mathbb{R}^n \to \mathbb{R}$ 是凸函数,且 f 在点 $\bar{\boldsymbol{x}} \in \mathbb{R}^n$ 处可微,若 $\nabla f(\bar{\boldsymbol{x}}) = \boldsymbol{0}$,则 $\bar{\boldsymbol{x}}$ 为问题(3.1.1)的全局极小点.

证明 因为 f 是 $\mathbb{R}^n$ 上的凸函数,且在点 $\bar{\boldsymbol{x}}$ 处可微,所以由定理 2.3.4,有

$$f(\boldsymbol{x}) \geqslant f(\bar{\boldsymbol{x}}) + \nabla f(\bar{\boldsymbol{x}})^{\mathrm{T}}(\boldsymbol{x} - \bar{\boldsymbol{x}}), \forall \boldsymbol{x} \in \mathbb{R}^n.$$

由于 $\nabla f(\bar{\boldsymbol{x}}) = \boldsymbol{0}$,因此

$$f(\boldsymbol{x}) \geqslant f(\bar{\boldsymbol{x}}), \forall \boldsymbol{x} \in \mathbb{R}^n,$$

即 $\bar{\boldsymbol{x}}$ 为问题(3.1.1)的全局极小点. □

根据定理 3.1.4 和定理 3.1.1,立即得到

推论 3.1.5 设 $f:\mathbb{R}^n \to \mathbb{R}$ 是凸函数,且 f 在点 $\bar{\boldsymbol{x}} \in \mathbb{R}^n$ 处可微,则 $\bar{\boldsymbol{x}}$ 是问题(3.1.1)的全局极小点的充要条件是 $\nabla f(\bar{\boldsymbol{x}}) = \boldsymbol{0}$. □

例 3.1.4 试证正定二次函数

$$f(\boldsymbol{x}) = \frac{1}{2}\boldsymbol{x}^{\mathrm{T}}\boldsymbol{Q}\boldsymbol{x} + \boldsymbol{b}^{\mathrm{T}}\boldsymbol{x} + c$$

有惟一的严格全局极小点

$$\bar{\boldsymbol{x}} = -\boldsymbol{Q}^{-1}\boldsymbol{b},$$

其中 $\boldsymbol{Q}$ 为 n 阶正定矩阵.

证明 因为 $\boldsymbol{Q}$ 为正定矩阵,且

$$\nabla f(\boldsymbol{x}) = \boldsymbol{Q}\boldsymbol{x} + \boldsymbol{b}, \forall \boldsymbol{x} \in \mathbb{R}^n.$$

所以得 $f(\boldsymbol{x})$ 的惟一平稳点 $\bar{\boldsymbol{x}} = -\boldsymbol{Q}^{-1}\boldsymbol{b}$. 又由于 f 是严格凸函数,因此由定理 3.1.4 知,$\bar{\boldsymbol{x}}$ 是 $f(\boldsymbol{x})$ 的严格全局极小点. □

3.2 等式约束最优化问题的最优性条件

考虑等式约束最优化问题

$$\begin{cases} \min & f(\boldsymbol{x}); \\ \text{s.t.} & h_j(\boldsymbol{x}) = 0, j = 1,2,\cdots,l, \end{cases} \tag{3.2.1}$$

其中 $f:\mathbb{R}^n \to \mathbb{R}, h_j:\mathbb{R}^n \to \mathbb{R}(j=1,2,\cdots,l)$. 这是经典的条件极值问题,它的最优性条件和求解方法在微积分学中已从理论上得到解决,这里加以总结和推广.

定理 3.2.1 设 $f:\mathbb{R}^n \to \mathbb{R}$ 在点 $\bar{\boldsymbol{x}} \in \mathbb{R}^n$ 处可微,$h_j:\mathbb{R}^n \to \mathbb{R}(j=1,2,\cdots,l)$ 在点 $\bar{\boldsymbol{x}}$ 处具有一阶连续偏导数,向量组 $\nabla h_1(\bar{\boldsymbol{x}}), \nabla h_2(\bar{\boldsymbol{x}}), \cdots, \nabla h_l(\bar{\boldsymbol{x}})$ 线性无关. 若 $\bar{\boldsymbol{x}}$ 是问题(3.2.1)的局部极小点,则存在实数 $\bar{v}_j, j=1,2,\cdots,l$,使得

$$\nabla f(\bar{\boldsymbol{x}}) - \sum_{j=1}^{l} \bar{v}_j \nabla h_j(\bar{\boldsymbol{x}}) = \boldsymbol{0}. \tag{3.2.2}$$

这就是 Lagrange 定理,式(3.2.2)就是问题(3.2.1)局部极小点的一阶必要条件.

在证明这个定理之前,先分析它的意义.

我们称 $n+l$ 元函数

$$L(\boldsymbol{x},\boldsymbol{v}) = f(\boldsymbol{x}) - \boldsymbol{v}^{\mathrm{T}}\boldsymbol{h}(\boldsymbol{x})$$

为 Lagrange 函数(Lagrange function),其中 $\boldsymbol{v} = (v_1, v_2, \cdots, v_l)^{\mathrm{T}}$,称之为 Lagrange 乘子向量(Lagrange multipliers vector),$\boldsymbol{h}(\boldsymbol{x}) = (h_1(\boldsymbol{x}), h_2(\boldsymbol{x}), \cdots, h_l(\boldsymbol{x}))^{\mathrm{T}}$,则

$$\nabla L(\boldsymbol{x},\boldsymbol{v}) = \begin{bmatrix} \nabla_x L \\ \nabla_v L \end{bmatrix},$$

这里

$$\nabla_x L(\boldsymbol{x},\boldsymbol{v}) = \nabla f(\boldsymbol{x}) - \sum_{j=1}^{l} v_j \nabla h_j(\boldsymbol{x}),$$

$$\nabla_v L(\boldsymbol{x},\boldsymbol{v}) = -\boldsymbol{h}(\boldsymbol{x}).$$

由此可知，无约束最优化问题

$$\min L(\boldsymbol{x},\boldsymbol{v}) \tag{3.2.3}$$

的最优性条件

$$\nabla L(\bar{\boldsymbol{x}},\bar{\boldsymbol{v}}) = \boldsymbol{0}$$

恰好给出了约束最优化问题(3.2.1)的最优性条件(3.2.2)和等式约束

$$h_j(\bar{\boldsymbol{x}}) = 0, j = 1,2,\cdots,l,$$

此即问题(3.2.1)的局部极小点 $\bar{\boldsymbol{x}}$ 必须满足的约束条件.

于是，从方程组

$$\begin{cases}\nabla f(\boldsymbol{x}) - \sum_{j=1}^{l} v_j \nabla h_j(\boldsymbol{x}) = \boldsymbol{0},\\ h_j(\boldsymbol{x}) = 0, j = 1,2,\cdots,l\end{cases}$$

解出 $L(\boldsymbol{x},\boldsymbol{v})$ 的平稳点 $\begin{bmatrix}\bar{\boldsymbol{x}}\\ \bar{\boldsymbol{v}}\end{bmatrix}$，那么对应的 $\bar{\boldsymbol{x}}$ 就可能是问题(3.2.1)的最优解. 这表明，Lagrage 定理的意义在于能将等式约束最优化问题(3.2.1)的求解化为无约束最优化问题(3.2.3)的求解.

定理 3.2.1 的证明 因为 $\nabla h_1(\bar{\boldsymbol{x}}),\nabla h_2(\bar{\boldsymbol{x}}),\cdots,\nabla h_l(\bar{\boldsymbol{x}})$ 线性无关，所以向量函数 $\boldsymbol{h}(\boldsymbol{x})$ 在 $\bar{\boldsymbol{x}}$ 处的 Jacobi 矩阵

$$\nabla \boldsymbol{h}(\bar{\boldsymbol{x}}) = (\nabla h_1(\bar{\boldsymbol{x}}),\nabla h_2(\bar{\boldsymbol{x}}),\cdots,\nabla h_l(\bar{\boldsymbol{x}}))^{\mathrm{T}}$$

是行满秩的，不妨设 $\nabla \boldsymbol{h}(\bar{\boldsymbol{x}})$ 的前 l 列是线性无关的，从而线性方程组

$$\sum_{j=1}^{l} \frac{\partial h_j(\bar{\boldsymbol{x}})}{\partial x_i} v_j = \frac{\partial f(\bar{\boldsymbol{x}})}{\partial x_i}, i = 1,2,\cdots,l \tag{3.2.4}$$

有惟一的解 $\bar{\boldsymbol{v}} = (\bar{v}_1,\bar{v}_2,\cdots,\bar{v}_l)^{\mathrm{T}}$.

记 $\boldsymbol{y} = (x_{l+1},x_{l+2},\cdots,x_n)^{\mathrm{T}}$，则对于等式约束方程

$$h_j(\boldsymbol{x}) = h_j(x_1,x_2,\cdots,x_l,\boldsymbol{y}^{\mathrm{T}}) = 0, j = 1,2,\cdots,l,$$

由隐函数存在定理可知，在 $\bar{\boldsymbol{y}} = (\bar{x}_{l+1},\bar{x}_{l+2},\cdots,\bar{x}_n)^{\mathrm{T}}$ 的某邻域内存在 $n-l$ 元函数 $\psi_i, i = 1,2,\cdots,l$，使

$$x_i = \psi_i(\boldsymbol{y}), i = 1,2,\cdots,l,$$

并且

$$h_j(\psi_1(\boldsymbol{y}),\psi_2(\boldsymbol{y}),\cdots,\psi_l(\boldsymbol{y}),\boldsymbol{y}^{\mathrm{T}}) = 0, j = 1,2,\cdots,l. \tag{3.2.5}$$

另外，由于 $\bar{\boldsymbol{x}}$ 是问题(3.2.1)的局部极小点，因此存在 $\delta > 0$，使

$$f(\bar{\boldsymbol{x}}) \leqslant f(\boldsymbol{x}), \forall \boldsymbol{x} \in N_\delta(\bar{\boldsymbol{x}}) \cap \{\boldsymbol{x} \mid \boldsymbol{h}(\boldsymbol{x}) = \boldsymbol{0}\}.$$

从而对于 $\bar{\boldsymbol{y}}$ 的某邻域内的 $\boldsymbol{y}$，有

$$f(\psi_1(\bar{\boldsymbol{y}}),\psi_2(\bar{\boldsymbol{y}}),\cdots,\psi_l(\bar{\boldsymbol{y}}),\bar{\boldsymbol{y}}^{\mathrm{T}}) \leqslant f(\psi_1(\boldsymbol{y}),\psi_2(\boldsymbol{y}),\cdots,\psi_l(\boldsymbol{y}),\boldsymbol{y}^{\mathrm{T}}),$$

即 $\bar{\boldsymbol{y}}$ 是无约束最优化问题

$$\min \varphi(\boldsymbol{y}) = f(\psi_1(\boldsymbol{y}),\psi_2(\boldsymbol{y}),\cdots,\psi_l(\boldsymbol{y}),\boldsymbol{y}^{\mathrm{T}})$$

的局部极小点. 于是由定理 3.1.1 知

$$\nabla \varphi(\bar{\boldsymbol{y}}) = \boldsymbol{0},$$

即有

$$\frac{\partial \varphi(\bar{\boldsymbol{y}})}{\partial x_i} = \sum_{k=1}^{l} \frac{\partial f(\bar{\boldsymbol{x}})}{\partial x_k} \frac{\partial \psi_k(\bar{\boldsymbol{y}})}{\partial x_i} + \frac{\partial f(\bar{\boldsymbol{x}})}{\partial x_i} = 0, i = l+1, l+2, \cdots, n. \tag{3.2.6}$$

对(3.2.5)式应用隐函数求导法则,得

$$\sum_{k=1}^{l} \frac{\partial h_j(\bar{\boldsymbol{x}})}{\partial x_k} \frac{\partial \psi_k(\bar{\boldsymbol{y}})}{\partial x_i} + \frac{\partial h_j(\bar{\boldsymbol{x}})}{\partial x_i} = 0, j = 1, 2, \cdots, l; i = l+1, l+2, \cdots, n.$$

把上式各方程相应地乘以 $\bar{v}_j$ 并相加,得

$$\sum_{j=1}^{l} \sum_{k=1}^{l} \bar{v}_j \frac{\partial h_j(\bar{\boldsymbol{x}})}{\partial x_k} \frac{\partial \psi_k(\bar{\boldsymbol{y}})}{\partial x_i} + \sum_{j=1}^{l} \bar{v}_j \frac{\partial h_j(\bar{\boldsymbol{x}})}{\partial x_i} = 0, i = l+1, l+2, \cdots, n. \tag{3.2.7}$$

将(3.2.6)式减去(3.2.7)式,有

$$\sum_{k=1}^{l} \left(\frac{\partial f(\bar{\boldsymbol{x}})}{\partial x_k} - \sum_{j=1}^{l} \bar{v}_j \frac{\partial h_j(\bar{\boldsymbol{x}})}{\partial x_k} \right) \frac{\partial \psi_k(\bar{\boldsymbol{y}})}{\partial x_i} + \frac{\partial f(\bar{\boldsymbol{x}})}{\partial x_i} - \sum_{j=1}^{l} \bar{v}_j \frac{\partial h_j(\bar{\boldsymbol{x}})}{\partial x_i} = 0,$$

$$i = l+1, l+2, \cdots, n.$$

由(3.2.4)式知,上式括弧中的式子为 0,从而

$$\frac{\partial f(\bar{\boldsymbol{x}})}{\partial x_i} - \sum_{j=1}^{l} \bar{v}_j \frac{\partial h_j(\bar{\boldsymbol{x}})}{\partial x_i} = 0, i = l+1, l+2, \cdots, n. \tag{3.2.8}$$

综合(3.2.4)式和(3.2.8)式,即得(3.2.2)式. □

下面给出问题(3.2.1)的局部极小点的二阶充分条件.

定理 3.2.2 设 $f: \mathbb{R}^n \to \mathbb{R}$ 和 $h_j: \mathbb{R}^n \to \mathbb{R} (j = 1, 2, \cdots, l)$ 在点 $\bar{\boldsymbol{x}} \in \mathbb{R}^n$ 的处具有二阶连续偏导数. 若存在 $\bar{\boldsymbol{v}} \in \mathbb{R}^l$,使得

$$\nabla L(\bar{\boldsymbol{x}}, \bar{\boldsymbol{v}}) = \boldsymbol{0},$$

并且,对一切非零向量 $\boldsymbol{z} \in \mathbb{R}^n$,只要

$$\boldsymbol{z}^{\mathrm{T}} \nabla h_j(\bar{\boldsymbol{x}}) = 0, j = 1, 2, \cdots, l,$$

便有

$$\boldsymbol{z}^{\mathrm{T}} \nabla_{xx}^2 L(\bar{\boldsymbol{x}}, \bar{\boldsymbol{v}}) \boldsymbol{z} > 0,$$

则 $\bar{\boldsymbol{x}}$ 是问题(3.2.1)的严格局部极小点.

证明 由 $\nabla L(\bar{\boldsymbol{x}}, \bar{\boldsymbol{v}}) = \boldsymbol{0}$ 可知 $\nabla_v L(\bar{\boldsymbol{x}}, \bar{\boldsymbol{v}}) = 0$,即 $\boldsymbol{h}(\bar{\boldsymbol{x}}) = 0$,因此 $\bar{\boldsymbol{x}}$ 是问题(3.2.1)的可行点.

用反证法证明 $\bar{\boldsymbol{x}}$ 是问题(3.2.1)的严格局部极小点,若不然,在 $\bar{\boldsymbol{x}}$ 的某个邻域内存在收敛于 $\bar{\boldsymbol{x}}$ 的点列 $\{\boldsymbol{y}_k\}$, $\boldsymbol{y}_k \neq \bar{\boldsymbol{x}}, k = 1, 2, \cdots$,使得对每个 $\boldsymbol{y}_k$,有

$$h_j(\boldsymbol{y}_k) = 0, j = 1, 2, \cdots, l, \tag{3.2.9}$$

$$f(\boldsymbol{y}_k)\leqslant f(\bar{\boldsymbol{x}}). \tag{3.2.10}$$

令 $\theta_k = \|\boldsymbol{y}_k - \bar{\boldsymbol{x}}\|, \boldsymbol{z}_k = (\boldsymbol{y}_k - \bar{\boldsymbol{x}})/\theta_k$,则

$$\boldsymbol{y}_k = \bar{\boldsymbol{x}} + \theta_k \boldsymbol{z}_k, \theta_k > 0, \|\boldsymbol{z}_k\| = 1.$$

由于$\{\boldsymbol{z}_k\}$有界,因此有收敛子列,不妨设$\{\boldsymbol{z}_k\}$收敛,记 $\boldsymbol{z}_k \to \boldsymbol{z}(k\to\infty)$,则 $\|\boldsymbol{z}\| = 1$. 而 $\theta_k \to 0(k\to\infty)$,故由 Taylor 公式有

$$h_j(\boldsymbol{y}_k) = h_j(\bar{\boldsymbol{x}}) + \theta_k \boldsymbol{z}_k^{\mathrm{T}} \nabla h_j(\bar{\boldsymbol{x}}) + o(\|\theta_k \boldsymbol{z}_k\|), j = 1,2,\cdots,l.$$

注意到 $\bar{\boldsymbol{x}}$ 为可行点和(3.2.9)式,并在上式中除以 θ_k,再令 $k\to\infty$,可得

$$\boldsymbol{z}^{\mathrm{T}} \nabla h_j(\bar{\boldsymbol{x}}) = 0, j = 1,2,\cdots,l.$$

同样,由二阶 Taylor 公式,有

$$L(\boldsymbol{y}_k,\bar{\boldsymbol{v}}) = L(\bar{\boldsymbol{x}},\bar{\boldsymbol{v}}) + \theta_k \boldsymbol{z}_k^{\mathrm{T}} \nabla_x L(\bar{\boldsymbol{x}},\bar{\boldsymbol{v}}) + \frac{1}{2}\theta_k^2 \boldsymbol{z}_k^{\mathrm{T}} \nabla_{xx}^2 L(\bar{\boldsymbol{x}},\bar{\boldsymbol{v}})\boldsymbol{z}_k + o(\|\theta_k \boldsymbol{z}_k\|^2). \tag{3.2.11}$$

由于对于 $k=1,2,\cdots$和 $j=1,2,\cdots,l$,有 $h_j(\boldsymbol{y}_k)=0, h_j(\bar{\boldsymbol{x}})=0$,因此

$$L(\boldsymbol{y}_k,\bar{\boldsymbol{v}}) = f(\boldsymbol{y}_k) - \sum_{j=1}^{l} \bar{v}_j h_j(\boldsymbol{y}_k) = f(\boldsymbol{y}_k),$$

$$L(\bar{\boldsymbol{x}},\bar{\boldsymbol{v}}) = f(\bar{\boldsymbol{x}}) - \sum_{j=1}^{l} \bar{v}_j h_j(\bar{\boldsymbol{x}}) = f(\bar{\boldsymbol{x}}).$$

又因为$\nabla_x L(\bar{\boldsymbol{x}},\bar{\boldsymbol{v}}) = \boldsymbol{0}$,所以由(3.2.11)式得

$$f(\boldsymbol{y}_k) = f(\bar{\boldsymbol{x}}) + \frac{1}{2}\theta_k^2 \boldsymbol{z}_k^{\mathrm{T}} \nabla_{xx}^2 L(\bar{\boldsymbol{x}},\bar{\boldsymbol{v}})\boldsymbol{z}_k + o(\|\theta_k \boldsymbol{z}_k\|^2),$$

从而由(3.2.10)式得

$$\frac{1}{2}\theta_k^2 \boldsymbol{z}_k^{\mathrm{T}} \nabla_{xx}^2 L(\bar{\boldsymbol{x}},\bar{\boldsymbol{v}})\boldsymbol{z}_k + o(\|\theta_k \boldsymbol{z}_k\|^2) \leqslant 0.$$

对上式除以$\frac{1}{2}\theta_k^2$,并令 $k\to\infty$,即得

$$\boldsymbol{z}^{\mathrm{T}} \nabla_{xx}^2 L(\bar{\boldsymbol{x}},\bar{\boldsymbol{v}})\boldsymbol{z} \leqslant 0.$$

此与定理的条件矛盾. □

这个定理的几何意义是,在 Lagrange 函数的平稳点$\begin{bmatrix}\bar{\boldsymbol{x}}\\ \bar{\boldsymbol{v}}\end{bmatrix}$处,如果 Lagrange 函数关于 $\boldsymbol{x}$ 的 Hesse 矩阵在约束曲面的切平面上正定(并不需要在$\mathbb{R}^n$ 中正定),那么 $\bar{\boldsymbol{x}}$ 就是问题(3.2.1)的严格局部极小点.

例 3.2.1 试用最优性条件求解

$$\begin{cases}\min & f(\boldsymbol{x}) = x_1^2 + x_2^2;\\ \text{s.t.} & h(\boldsymbol{x}) = x_1 x_2 - 8 = 0.\end{cases}$$

解 Lagrange 函数为

$$L(\boldsymbol{x},v) = x_1^2 + x_2^2 - v(x_1 x_2 - 8),$$

则

$$\nabla L(\boldsymbol{x},v)=\begin{bmatrix}2x_1-vx_2\\2x_2-vx_1\\-(x_1x_2-8)\end{bmatrix},$$

从而得 $L(\boldsymbol{x},v)$ 的平稳点 $(\sqrt{8},\sqrt{8},2)^{\mathrm{T}}$ 和 $(-\sqrt{8},-\sqrt{8},2)^{\mathrm{T}}$. 对应有

$$\bar{\boldsymbol{x}}=(\sqrt{8},\sqrt{8})^{\mathrm{T}},\bar{v}=2,$$

和

$$\bar{\boldsymbol{x}}=(-\sqrt{8},-\sqrt{8})^{\mathrm{T}},\bar{v}=2.$$

由于

$$\nabla_{xx}^2L(\bar{\boldsymbol{x}},\bar{v})=\begin{bmatrix}2&-\bar{v}\\-\bar{v}&2\end{bmatrix}=\begin{bmatrix}2&-2\\-2&2\end{bmatrix},\nabla h(\bar{\boldsymbol{x}})=\begin{bmatrix}\bar{x}_2\\\bar{x}_1\end{bmatrix},$$

因此

$$\begin{aligned}M(\bar{\boldsymbol{x}})&=\{(z_1,z_2)^{\mathrm{T}}\mid(z_1,z_2)\nabla h(\bar{\boldsymbol{x}})=0\}\\&=\{(z_1,z_2)^{\mathrm{T}}\mid z_1\bar{x}_2+z_2\bar{x}_1=0\}\\&=\{(z_1,z_2)^{\mathrm{T}}\mid z_1=-z_2\},\end{aligned}$$

并且对任何 $z\in M(\bar{\boldsymbol{x}}),z\neq\boldsymbol{0}$,有

$$z^{\mathrm{T}}\nabla_{xx}^2L(\bar{\boldsymbol{x}},\bar{v})z=2z_1^2-4z_1z_2+2z_2^2=8z_1^2>0.$$

利用定理 3.2.2,所得的两个可行点 $\bar{\boldsymbol{x}}=(\sqrt{8},\sqrt{8})^{\mathrm{T}}$ 和 $\bar{\boldsymbol{x}}=(-\sqrt{8},-\sqrt{8})^{\mathrm{T}}$ 都是问题的严格局部极小点. □

3.3 不等式约束最优化问题的最优性条件

考虑不等式约束最优化问题

$$\begin{cases}\min & f(\boldsymbol{x});\\\text{s.t.} & g_i(\boldsymbol{x})\geqslant0,i=1,2,\cdots,m,\end{cases}\tag{3.3.1}$$

其中 $f:\mathbb{R}^n\to\mathbb{R},g_i:\mathbb{R}^n\to\mathbb{R}(i=1,2,\cdots,m)$.

3.3.1 几何最优性条件

为增加直观性,我们先给出问题(3.3.1)最优性条件的几何表示,然后再讨论它的代数表示. 为此引入几个概念.

$\mathbb{R}^n$ 中集合 S 的闭包(closure)是指

$$\mathrm{cl}S=\{x\in\mathbb{R}^n\mid S\cap N_\delta(\boldsymbol{x})\neq\varnothing,\forall\delta>0\}.$$

定义 3.3.1 设 $S\subseteq\mathbb{R}^n,\bar{\boldsymbol{x}}\in\mathrm{cl}S,\boldsymbol{d}\in\mathbb{R}^n,\boldsymbol{d}\neq\boldsymbol{0}$,若存在 $\delta>0$,使得

$$\bar{x}+\lambda d\in S,\forall\lambda\in(0,\delta),$$

则称 d 为集合 S 在点 $\bar{x}$ 处的可行方向(feasible direction).

如果 d 既是 S 在点 $\bar{x}$ 处的可行方向,又是 $f:S\to\mathbb{R}$ 在点 $\bar{x}$ 处的下降方向,则称 d 为 f 在点 $\bar{x}$ 处关于 S 的可行下降方向(feasible descent direction).

由此定义可知,从点 $\bar{x}$ 出发,沿可行方向 d 的微小移动,必能达到 S 的一个点.

集合 S 在点 $\bar{x}$ 处所有可行方向的集合

$$D=\{d\mid d\neq 0,\text{存在 }\delta>0,\text{使 }\bar{x}+\lambda d\in S,\forall\lambda\in(0,\delta)\},$$

显然是一个以 $\bar{x}$ 为顶点的锥,我们称之为 S 在 $\bar{x}$ 处的可行方向锥(feasible direction cone).

特别地,当 $\bar{x}\in \text{int}S$ 时,S 在点 $\bar{x}$ 处的可行方向锥是全空间 $\mathbb{R}^n$.

根据定理 1.2.3,如果 $S\subseteq\mathbb{R}^n,\bar{x}\in S,d\in\mathbb{R}^n$ 且 $f:S\to\mathbb{R}$ 在点 $\bar{x}$ 处可微,$\nabla f(\bar{x})^{\mathrm{T}}d<0$,则 d 为 f 在点 $\bar{x}$ 处的下降方向. 因此,我们称锥

$$F_0=\{d\mid\nabla f(\bar{x})^{\mathrm{T}}d<0\}$$

为 f 在点 $\bar{x}$ 处的下降方向锥(descent direction cone).

图 3.3.1 画出了可行方向锥和下降方向锥,从几何图形上看,在点 $\bar{x}$ 处沿下降方向移动,就导致目标函数 f 的值减少. 所以,在极小点处,任何下降方向都不是可行方向,而任何可行方向也不是下降方向,就是说,不存在可行下降方向. 这样就得到了一般约束最优化问题极小点的一个必要条件.

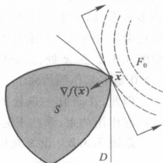


图 3.3.1 可行方向锥与下降方向锥

定理 3.3.1 考虑约束最优化问题

$$\min_{x\in S}f(x),\tag{3.3.2}$$

其中 $S\subseteq\mathbb{R}^n$ 是非空集合,$f:S\to\mathbb{R}$. 设 $\bar{x}\in S$,且 f 在 $\bar{x}$ 处可微. 若 $\bar{x}$ 是问题(3.3.2)的局部极小点,则

$$F_0\cap D=\varnothing.\tag{3.3.3}$$

证明 用反证法. 假若存在 $d\in F_0\cap D$,则 $d\in F_0,d\in D,d\neq 0$. 由 F_0 的定义和定理 1.2.3,存在 $\delta_1>0$,使

$$f(\bar{x}+\lambda d)<f(\bar{x}),\forall\lambda\in(0,\delta_1);$$

由 D 的定义,存在 $\delta_2>0$,使

$$\bar{x}+\lambda d\in S,\forall\lambda\in(0,\delta_2);$$

再由 $\bar{x}$ 为局部极小点的定义,存在 $\delta_3>0$,使

$$f(x)\geqslant f(\bar{x}),\forall x\in N_{\delta_3}(\bar{x})\cap S.\tag{3.3.4}$$

取 $\delta = \min\left\{\delta_1, \delta_2, \frac{\delta_3}{\|\boldsymbol{d}\|}\right\}$，则 $\bar{\boldsymbol{x}} + \frac{\delta}{2}\boldsymbol{d} \in N_{\delta_3}(\bar{\boldsymbol{x}}) \cap S$ 且

$$f\left(\bar{\boldsymbol{x}} + \frac{\delta}{2}\boldsymbol{d}\right) < f(\bar{\boldsymbol{x}}),$$

此与(3.3.4)式矛盾. □

现在再回头来讨论问题(3.3.1)的最优性条件.

设 $\bar{\boldsymbol{x}}$ 为问题(3.3.1)的可行点，则其不等式约束条件在 $\bar{\boldsymbol{x}}$ 处呈现出两种情形：

(1) $g_i(\bar{\boldsymbol{x}}) = 0$，称第 i 个不等式约束为在 $\bar{\boldsymbol{x}}$ 处起作用的约束(active constraint)或紧约束(tight constraint)；

(2) $g_i(\bar{\boldsymbol{x}}) > 0$，称第 i 个不等式约束为在 $\bar{\boldsymbol{x}}$ 处不起作用的约束(inactive constraint)或松约束(loose constraint)；

我们用 $I(\bar{\boldsymbol{x}})$ 表示在可行点 $\bar{\boldsymbol{x}}$ 处起作用约束的指标集，即

$$I(\bar{\boldsymbol{x}}) = \{i \mid g_i(\bar{\boldsymbol{x}}) = 0, i = 1,2,\cdots,m\}.$$

对于起作用约束，当点沿某些方向稍微离开 $\bar{\boldsymbol{x}}$ 时，仍能满足这些约束，而沿另一些方向离开 $\bar{\boldsymbol{x}}$ 时，不论步长多么小，都会违背这些约束. 对于不起作用约束，当点稍微离开 $\bar{\boldsymbol{x}}$ 时，不论什么方向都不违背这些约束. 如图 3.3.2，在 $\bar{\boldsymbol{x}}$ 处，$g_1 \geq 0$ 和 $g_2 \geq 0$ 是起作用约束，$g_3 \geq 0$ 是不起作用约束.

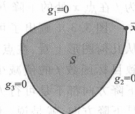


图 3.3.2 起作用约束与不起作用约束

因此，在研究一点处的可行方向时，只需考虑在该点起作用约束，那些不起作用约束可以暂且不管.

定理 3.3.2 设 $\bar{\boldsymbol{x}}$ 是问题(3.3.1)的可行点，f 和 $g_i(i \in I(\bar{\boldsymbol{x}}))$ 在点 $\bar{\boldsymbol{x}}$ 处可微，$g_i(i \notin I(\bar{\boldsymbol{x}}))$ 在点 $\bar{\boldsymbol{x}}$ 处连续，如果 $\bar{\boldsymbol{x}}$ 是问题(3.3.1)的局部极小点，则

$$F_0 \cap G_0 = \varnothing, \tag{3.3.5}$$

其中

$$G_0 = \{\boldsymbol{d} \mid \nabla g_i(\bar{\boldsymbol{x}})^{\mathrm{T}}\boldsymbol{d} > 0, i \in I(\bar{\boldsymbol{x}})\}.$$

证明 由定理 3.3.1 知，只需证明 $G_0 \subseteq D$.

对于任意 $\boldsymbol{d} \in G_0$，当 $i \in I(\bar{\boldsymbol{x}})$ 时，$\nabla g_i(\bar{\boldsymbol{x}})^{\mathrm{T}}\boldsymbol{d} > 0$，所以由定理 1.2.3，$\boldsymbol{d}$ 是 $g_i(\boldsymbol{x})$ 在点 $\bar{\boldsymbol{x}}$ 处的上升方向，即存在 $\delta_i > 0$，使得对一切 $\lambda \in (0, \delta_i)$，有

$$g_i(\bar{\boldsymbol{x}} + \lambda\boldsymbol{d}) > g_i(\bar{\boldsymbol{x}}) = 0, i \in I(\bar{\boldsymbol{x}}).$$

而当 $i \notin I(\bar{\boldsymbol{x}})$ 时，因为 $g_i(\bar{\boldsymbol{x}}) > 0$，所以由 $g_i(\boldsymbol{x})$ 在点 $\bar{\boldsymbol{x}}$ 处连续知，存在 $\delta_i > 0$，使得对一切 $\lambda \in (0, \delta_i)$，有

$$g_i(\bar{\boldsymbol{x}} + \lambda\boldsymbol{d}) > 0, i \notin I(\bar{\boldsymbol{x}}).$$

取 $\delta=\min\{\delta_i \mid i=1,2,\cdots,m\}$，则对一切 $\lambda\in(0,\delta)$，有

$$g_i(\bar{x}+\lambda d)>0, i=1,2,\cdots,m,$$

从而 $\bar{x}+\lambda d$ 为问题(3.3.1)的可行点. 显然 $d\neq 0$，因此 d 是在点 $\bar{x}$ 处的一个可行方向，即 $d\in D$，从而 $G_0\subseteq D$. □

条件(3.3.3)和(3.3.5)分别称问题(3.3.2)和(3.3.1)的几何最优性条件(geometric conditions for optimality)，其特点是直观，但难以在实际计算中应用. 下面我们将几何最优性条件转化为代数最优性条件(algebraic conditions for optimality).

3.3.2 Fritz John 条件

定理 3.3.3 设 $\bar{x}$ 是问题(3.3.1)的可行点，f 和 $g_i(i\in I(\bar{x}))$ 在点 $\bar{x}$ 处可微，$g_i(i\notin I(\bar{x}))$ 在点 $\bar{x}$ 处连续，若 $\bar{x}$ 是问题(3.3.1)的局部极小点，则存在不全为 0 的非负数 $u_0,u_i(i\in I(\bar{x}))$，使

$$u_0\nabla f(\bar{x})-\sum_{i\in I(\bar{x})}u_i\nabla g_i(\bar{x})=\mathbf{0}. \tag{3.3.6}$$

如果 $g_i(i\notin I(\bar{x}))$ 在点 $\bar{x}$ 处也可微，则存在不全为 0 的非负数 $u_0,u_1,\cdots,u_m$，使

$$\begin{cases}u_0\nabla f(\bar{x})-\sum_{i=1}^{m}u_i\nabla g_i(\bar{x})=\mathbf{0},\\ u_ig_i(\bar{x})=0, i=1,2,\cdots,m.\end{cases} \tag{3.3.7}$$

证明 根据定理 3.3.2，$F_0\cap G_0=\varnothing$，即关系式组

$$\begin{cases}\nabla f(\bar{x})^{\mathrm{T}}d<0,\\ -\nabla g_i(\bar{x})^{\mathrm{T}}d<0, i\in I(\bar{x})\end{cases}$$

无解. 又由 Gordan 定理，必存在不全为 0 的非负数 $u_0,u_i(i\in I(\bar{x}))$，使(3.3.6)式成立.

如果 $g_i(i\notin I(\bar{x}))$ 在点 $\bar{x}$ 处可微，则只要 $u_i=0(i\notin I(\bar{x}))$，就有(3.3.7)式成立. □

式(3.3.6)和式(3.3.7)都称为问题(3.3.1)的 Fritz John 条件. 满足 Fritz John 条件的点称为 Fritz John 点.

例 3.3.1 设

$$\begin{cases}\min & f(x)=-x_1;\\ \text{s.t.} & g_1(x)=(1-x_1)^3-x_2\geqslant 0,\\ & g_2(x)=x_2\geqslant 0,\end{cases}$$

试判别 $\bar{x}=(1,0)^{\mathrm{T}}$ 是否为 Fritz John 点.

解 因为

$$\nabla f(\bar{x}) = \begin{bmatrix} -1 \\ 0 \end{bmatrix}, \nabla g_1(\bar{x}) = \begin{bmatrix} 0 \\ -1 \end{bmatrix}, \nabla g_2(\bar{x}) = \begin{bmatrix} 0 \\ 1 \end{bmatrix},$$

且 $I(\bar{x}) = \{1,2\}$,所以为使 Fritz John 条件

$$u_0\begin{bmatrix} -1 \\ 0 \end{bmatrix} - u_1\begin{bmatrix} 0 \\ -1 \end{bmatrix} - u_2\begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

成立,只有 $u_0 = 0$ 才行. 取 $u_0 = 0, u_1 = u_2 = \alpha > 0$ 即可,因此 $\bar{x}$ 是 Fritz John 点. □

这个不等式约束最优化问题的可行域见图 3.3.3,由此即知 $\bar{x} = (1,0)^{\mathrm{T}}$ 是局部极小点,也是全局极小点.

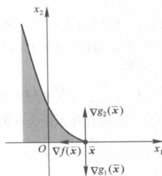


图 3.3.3 Fritz John 条件中 $u_0 = 0$ 的情形

这个例子说明在 Fritz John 条件中有可能 $u_0 = 0$,当 $u_0 = 0$ 时,目标函数的梯度 $\nabla f(\bar{x})$ 就会从 Fritz John 条件中消失,此时,Fritz John 条件实际上不包含目标函数的任何信息,仅仅表明起作用约束函数的梯度线性相关,而这对表述最优点没有什么实际价值.

为了保证 $u_0 > 0$,还需要对约束再加上一些限制条件. 这种限制条件通常称为约束规格(constraint qualifications). 一个自然的想法是附加 $\nabla g_i(\bar{x})$ $(i \in I(\bar{x}))$ 线性无关的约束规格(当然还有许多其他的约束规格,参见文献[48, 50]),这样就得到了著名的 Kuhn - Tucker 条件.

3.3.3 Kuhn - Tucker 条件

定理 3.3.4 设 $\bar{x}$ 是问题(3.3.1)的可行点,f 和 $g_i(i \in I(\bar{x}))$ 在点 $\bar{x}$ 处可微,$g_i(i \notin I(\bar{x}))$ 在点 $\bar{x}$ 处连续,并且 $\nabla g_i(\bar{x})$ $(i \in I(\bar{x}))$ 线性无关. 若 $\bar{x}$ 是问题(3.3.1)的局部极小点,则存在非负数 $u_i(i \in I(\bar{x}))$,使得

$$\nabla f(\bar{x}) - \sum_{i \in I(\bar{x})} u_i \nabla g_i(\bar{x}) = \mathbf{0}. \tag{3.3.8}$$

如果 $g_i(i \notin I(\bar{x}))$ 在点 $\bar{x}$ 处也可微，则存在 $u_i \geqslant 0(i=1,2,\cdots,m)$，使得

$$\begin{cases}\nabla f(\bar{\boldsymbol{x}}) - \sum_{i=1}^{m} u_i \nabla g_i(\bar{\boldsymbol{x}}) = \boldsymbol{0}, \\ u_i g_i(\bar{\boldsymbol{x}}) = 0, i = 1,2,\cdots,m.\end{cases} \tag{3.3.9}$$

证明　根据定理 3.3.3，存在不全为 0 的非负数 $\bar{u}_0, \bar{u}_i(i \in I(\bar{x}))$，使

$$\bar{u}_0 \nabla f(\bar{\boldsymbol{x}}) - \sum_{i \in I(\bar{x})} \bar{u}_i \nabla g_i(\bar{\boldsymbol{x}}) = \boldsymbol{0}.$$

易知上式中 $\bar{u}_0 > 0$，这是因为若 $\bar{u}_0 = 0$，则由 $\bar{u}_i(i \in I(\bar{x}))$ 不全为 0 知 $\nabla g_i(\bar{x})$ $(i \in I(\bar{x}))$ 线性相关，此与条件矛盾. 由此令

$$u_i = \frac{\bar{u}_i}{\bar{u}_0}, i \in I(\bar{\boldsymbol{x}}),$$

则 $u_i \geqslant 0(i \in I(\bar{x}))$，且(3.3.8)式成立.

如果 $g_i(i \notin I(\bar{x}))$ 在点 $\bar{x}$ 处也可微，则只要令 $u_i = 0(i \notin I(\bar{x}))$，就有(3.3.9)式成立. □

条件(3.3.8)和(3.3.9)均称为问题(3.3.1)的 Kuhn – Tucker 条件，简称 K – T条件. 满足 K – T 条件的点称为 K – T 点. 式(3.3.9)中的第 2 式称为互补松弛条件(complementary slackness condition).

K – T 条件有着明显的几何意义. 在 K – T 条件(3.3.8)中，

$$\nabla f(\bar{\boldsymbol{x}}) = \sum_{i \in I(\bar{x})} u_i \nabla g_i(\bar{\boldsymbol{x}}),$$

而集合

$$\left\{\boldsymbol{y} \mid \boldsymbol{y} = \sum_{i \in I(\bar{x})} u_i \nabla g_i(\bar{\boldsymbol{x}}), u_i \geqslant 0(i \in I(\bar{\boldsymbol{x}}))\right\}$$

是在点 $\bar{x}$ 处起作用约束函数的梯度 $\nabla g_i(\bar{x})(i \in I(\bar{x}))$ 所生成的凸锥. 因此，若 $\bar{x}$ 是问题(3.3.1)的局部极小点，则目标函数的梯度 $\nabla f(\bar{x})$ 必属于上述凸锥.

例 3.3.2　求最优化问题

$$\begin{cases}\min & f(\boldsymbol{x}) = (x_1 - 1)^2 + x_2; \\ \text{s.t.} & g_1(\boldsymbol{x}) = -x_1 - x_2 + 2 \geqslant 0, \\ & g_2(\boldsymbol{x}) = x_2 \geqslant 0\end{cases}$$

的 K – T 点.

解　因为

$$\nabla f(\boldsymbol{x}) = \begin{bmatrix} 2(x_1 - 1) \\ 1 \end{bmatrix}, \nabla g_1(\boldsymbol{x}) = \begin{bmatrix} -1 \\ -1 \end{bmatrix}, \nabla g_2(\boldsymbol{x}) = \begin{bmatrix} 0 \\ 1 \end{bmatrix},$$

所以 K – T 条件为

$$
\begin{cases}
2(x_1-1)+u_1=0,\\
1+u_1-u_2=0,\\
u_1(-x_1-x_2+2)=0,\\
u_2x_2=0,\\
u_1\geqslant 0, u_2\geqslant 0,
\end{cases}
\tag{3.3.10}
$$

这是以 x_1,x_2,u_1,u_2 为变元的非线性方程组，一般来说，求解非线性方程组比较复杂. 但是这里求解并不困难.

若 $u_2=0$，则由(3.3.10)的第 2 式得 $u_1=-1$，这与 $u_1\geqslant 0$ 矛盾. 因此 $u_2>0$，由(3.3.10)的第 4 式知 $x_2=0$；

若 $-x_1+2=0$，则由(3.3.10)的第 1 式得 $u_1=-2$，这与 $u_1\geqslant 0$ 矛盾. 因此由(3.3.10)式的第 3 式知 $u_1=0$；

再将 $u_1=0$ 代入(3.3.10)的第 1 式、第 2 式得 $x_1=1, u_2=1$.

由于 $u_1\geqslant 0$，$u_2\geqslant 0$，且 $\bar{x}=(1,0)^{\mathrm{T}}$ 为问题的可行点，因此 $\bar{x}$ 是问题的 K－T 点. □

下面给出凸规划的最优解的充分条件.

定理 3.3.5 设在问题(3.3.1)中，f 和 $-g_i(i=1,2,\cdots,m)$ 是凸函数，$\bar{x}$ 是可行点，并且 f 和 $g_i(i\in I(\bar{x}))$ 在点 $\bar{x}$ 处可微. 若 $\bar{x}$ 是问题(3.3.1)的 K－T 点，则 $\bar{x}$ 是问题(3.3.1)的全局极小点.

证明 由假设条件，问题(3.3.1)的可行域 S 是凸集. 因为 f 是凸函数且在点 $\bar{x}$ 处可微，所以

$$f(\boldsymbol{x})\geqslant f(\bar{\boldsymbol{x}})+\nabla f(\bar{\boldsymbol{x}})^{\mathrm{T}}(\boldsymbol{x}-\bar{\boldsymbol{x}}),\ \forall \boldsymbol{x}\in S.$$

又因 $\bar{x}$ 为 K－T 点，故存在 $u_i\geqslant 0(i\in I(\bar{x}))$，使

$$\nabla f(\bar{\boldsymbol{x}})=\sum_{i\in I(\bar{\boldsymbol{x}})}u_i\nabla g_i(\bar{\boldsymbol{x}}).$$

从而

$$f(\boldsymbol{x})\geqslant f(\bar{\boldsymbol{x}})+\sum_{i\in I(\bar{\boldsymbol{x}})}u_i\nabla g_i(\bar{\boldsymbol{x}})^{\mathrm{T}}(\boldsymbol{x}-\bar{\boldsymbol{x}}),\ \forall \boldsymbol{x}\in S. \tag{3.3.11}$$

同样，由于凸函数 $-g_i(i\in I(\bar{x}))$ 在点 $\bar{x}$ 处可微，因此对一切 $\boldsymbol{x}\in S$，有

$$-g_i(\boldsymbol{x})\geqslant -g_i(\bar{\boldsymbol{x}})+(-\nabla g_i(\bar{\boldsymbol{x}}))^{\mathrm{T}}(\boldsymbol{x}-\bar{\boldsymbol{x}}),\ \forall i\in I(\bar{\boldsymbol{x}}).$$

注意到上式中 $g_i(\bar{x})=0$，且对一切 $\boldsymbol{x}\in S$，有 $g_i(\boldsymbol{x})\geqslant 0$，从而

$$\nabla g_i(\bar{\boldsymbol{x}})^{\mathrm{T}}(\boldsymbol{x}-\bar{\boldsymbol{x}})\geqslant 0,\ \forall i\in I(\bar{\boldsymbol{x}}),\ \forall \boldsymbol{x}\in S, \tag{3.3.12}$$

根据(3.3.11)和(3.3.12)两式，有

$$f(\boldsymbol{x})\geqslant f(\bar{\boldsymbol{x}}),\ \forall \boldsymbol{x}\in S,$$

即 $\bar{x}$ 是问题(3.3.1)的全局极小点. □

由上述定理可知，例 3.3.2 中的 K－T 点 $\bar{x}=(1,0)^{\mathrm{T}}$ 必是全局极小点.

3.4 一般约束最优化问题的最优性条件

考虑等式和不等式约束最优化问题

$$\begin{cases}\min & f(\boldsymbol{x});\\ \text{s.t.} & g_i(\boldsymbol{x}) \geqslant 0, i = 1,2,\cdots,m,\\ & h_j(\boldsymbol{x}) = 0, j = 1,2,\cdots,l,\end{cases} \tag{3.4.1}$$

其中$f:\mathbb{R}^n \to \mathbb{R}, g_i:\mathbb{R}^n \to \mathbb{R}(i=1,2,\cdots,m), h_j:\mathbb{R}^n \to \mathbb{R}(j=1,2,\cdots,l)$. 并把问题(3.4.1)的可行域记为$S$.

同不等式约束最优化问题类似,我们先给出问题(3.4.1)的几何最优性条件,然后给出代数最优性条件.

设$\bar{\boldsymbol{x}}$为问题(3.4.1)的可行点,即$\bar{\boldsymbol{x}} \in S$,记点$\bar{\boldsymbol{x}}$处的紧约束的指标集为$I(\bar{\boldsymbol{x}}) = \{i \mid g_i(\bar{\boldsymbol{x}}) = 0, i = 1,2,\cdots,m\}$. 类似于定理3.3.2,我们有

定理3.4.1(几何最优性条件) 设$\bar{\boldsymbol{x}}$为问题(3.4.1)的可行点,f和$g_i(i \in I(\bar{\boldsymbol{x}}))$在点$\bar{\boldsymbol{x}}$处可微,$h_j(j=1,2,\cdots,l)$在点$\bar{\boldsymbol{x}}$处具有一阶连续偏导数,$g_i(i \notin I(\bar{\boldsymbol{x}}))$在点$\bar{\boldsymbol{x}}$处连续,并且向量组$\nabla h_1(\bar{\boldsymbol{x}}), \nabla h_2(\bar{\boldsymbol{x}}), \cdots, \nabla h_l(\bar{\boldsymbol{x}})$线性无关. 若$\bar{\boldsymbol{x}}$是问题(3.4.1)的局部极小点,则

$$F_0 \cap G_0 \cap H_0 = \varnothing,$$

这里

$$F_0 = \{\boldsymbol{d} \mid \nabla f(\bar{\boldsymbol{x}})^{\mathrm{T}}\boldsymbol{d} < 0\},$$

$$G_0 = \{\boldsymbol{d} \mid \nabla g_i(\bar{\boldsymbol{x}})^{\mathrm{T}}\boldsymbol{d} > 0, i \in I(\bar{\boldsymbol{x}})\},$$

$$H_0 = \{\boldsymbol{d} \mid \nabla h_j(\bar{\boldsymbol{x}})^{\mathrm{T}}\boldsymbol{d} = 0, j = 1,2,\cdots,l\}.$$

证明 若$l = n$,则由$\nabla h_1(\bar{\boldsymbol{x}}), \nabla h_2(\bar{\boldsymbol{x}}), \cdots, \nabla h_l(\bar{\boldsymbol{x}})$线性无关可知,$H_0 = \{\boldsymbol{0}\}$,从而$G_0 \cap H_0 = \varnothing$,故结论成立. 下设$l < n$,且$G_0 \cap H_0 \neq \varnothing$,我们只需证明:对于一切$\boldsymbol{d} \in G_0 \cap H_0$,必有$\boldsymbol{d} \notin F_0$.

由于$\nabla h_1(\bar{\boldsymbol{x}}), \nabla h_2(\bar{\boldsymbol{x}}), \cdots, \nabla h_l(\bar{\boldsymbol{x}})$线性无关,因此对于由它们生成的$\mathbb{R}^n$的$l$维子空间,存在一个正交补子空间. 因为$\boldsymbol{d} \in G_0 \cap H_0$,所以$\boldsymbol{d} \neq \boldsymbol{0}$,且$\nabla h_j(\bar{\boldsymbol{x}})^{\mathrm{T}}\boldsymbol{d} = 0 (j = 1,2,\cdots,l)$,从而可设这个正交补子空间的正交基为$\boldsymbol{d}, \boldsymbol{d}_1, \cdots, \boldsymbol{d}_{n-l-1}$.

现在考虑方程组

$$\begin{cases} h_j(\boldsymbol{x}) = 0, j = 1,2,\cdots,l,\\ \boldsymbol{d}_i^{\mathrm{T}}(\boldsymbol{x} - \bar{\boldsymbol{x}}) = 0, i = 1,2,\cdots,n-l-1,\\ \boldsymbol{d}^{\mathrm{T}}(\boldsymbol{x} - \bar{\boldsymbol{x}}) - \theta = 0,\end{cases} \tag{3.4.2}$$

其中θ为实变量. 若记

$$h(x) = (h_1(x), h_2(x), \cdots, h_l(x))^{\mathrm{T}}, A = (d_1, d_2, \cdots, d_{n-l-1}),$$

$$\varphi(x,\theta) = \begin{bmatrix} h(x) \\ A^{\mathrm{T}}(x - \bar{x}) \\ d^{\mathrm{T}}(x - \bar{x}) - \theta \end{bmatrix},$$

则问题(3.4.2)等价于

$$\varphi(x,\theta) = \mathbf{0}. \tag{3.4.3}$$

显然,向量函数 $\varphi(x,\theta)$ 在点$\begin{bmatrix} \bar{x} \\ 0 \end{bmatrix}$处关于 x 的 Jacobi 矩阵为

$$\nabla_x \varphi(\bar{x},0) = \begin{bmatrix} \nabla h(\bar{x}) \\ A^{\mathrm{T}} \\ d^{\mathrm{T}} \end{bmatrix},$$

其中$\nabla h(\bar{x}) = (\nabla h_1(\bar{x}), \nabla h_2(\bar{x}), \cdots, \nabla h_l(\bar{x}))^{\mathrm{T}}$ 是向量函数 $h(x)$ 在 $\bar{x}$ 点处的 Jacobi 矩阵. 由前面的讨论及假设,不难知道,$\nabla_x \varphi(\bar{x},0)$ 的逆矩阵为

$$(\nabla h(\bar{x})^{\mathrm{T}}(\nabla h(\bar{x}) \nabla h(\bar{x})^{\mathrm{T}})^{-1}, A, d).$$

又因为问题(3.4.2)的左边的每个函数关于 x 和 θ 都有一阶连续偏导数,所以由隐函数存在定理,在 $\theta = 0$ 的某个邻域内存在具有一阶连续偏导数的向量函数 $x(\theta)$,使 $x(0) = \bar{x}$,且当 θ 充分小时,有 $\varphi(x(\theta),\theta) = \mathbf{0}$,从而 $h(x(\theta)) = \mathbf{0}$.

把(3.4.3)式的两边对 θ 求导,得

$$\nabla_\theta \varphi(x,\theta) + \nabla_x \varphi(x,\theta) \nabla x(\theta) = \mathbf{0}.$$

而

$$\nabla_\theta \varphi(x,\theta) = (0,0,\cdots,0,-1)^{\mathrm{T}},$$

故

$$\nabla x(0) = -(\nabla_x \varphi(\bar{x},0))^{-1} \nabla_\theta \varphi(\bar{x},0) = d.$$

由于对一切 $i \notin I(\bar{x})$,$g_i(x)$ 在点 $\bar{x}$ 处连续,且 $g_i(\bar{x}) > 0$,因此当 θ 充分小时,有

$$g_i(x(\theta)) > 0, \forall i \notin I(\bar{x});$$

又因 $d \in G_0$,故 $g_i(x(\theta))$ 在 $\theta = 0$ 处的导数为

$$\nabla g_i(\bar{x})^{\mathrm{T}} \nabla x(0) = \nabla g_i(\bar{x})^{\mathrm{T}} d > 0, \forall i \in I(\bar{x}),$$

即 $g_i(x(\theta))(i \in I(\bar{x}))$ 在 $\theta = 0$ 处严格单调增加,即当 θ 为充分小的正数时,有

$$g_i(x(\theta)) > g_i(x(0)) = 0, \forall i \in I(\bar{x}).$$

于是当 θ 为充分小的正数时,有

$$\begin{cases} g_i(x(\theta)) > 0, i = 1,2,\cdots,m, \\ h_j(x(\theta)) = 0, j = 1,2,\cdots,l, \end{cases}$$

即当 $\theta > 0$ 充分小时,$x(\theta) \in S$.

因为 $\bar{x}$ 为问题(3.4.1)的局部极小点,所以当 $\theta > 0$ 充分小时,有

$$f(\boldsymbol{x}(\theta)) \geqslant f(\bar{\boldsymbol{x}}). \tag{3.4.4}$$

由于$f(\boldsymbol{x})$在点$\bar{\boldsymbol{x}}$处可微，$\boldsymbol{x}(\theta)$在$\theta=0$处可微，因此由Taylor公式，有

$$f(\boldsymbol{x}(\theta)) = f(\boldsymbol{x}(0)) + \theta \nabla f(\bar{\boldsymbol{x}})^{\mathrm{T}}\boldsymbol{d} + o(\theta),$$

即由(3.4.4)式知

$$\nabla f(\bar{\boldsymbol{x}})^{\mathrm{T}}\boldsymbol{d} + \frac{o(\theta)}{\theta} \geqslant 0.$$

在上式中令$\theta \to 0$，得$\nabla f(\bar{\boldsymbol{x}})^{\mathrm{T}}\boldsymbol{d} \geqslant 0$，即$\boldsymbol{d} \notin F_0$. □

下面给出这个几何最优性条件的代数表达.

定理 3.4.2(Fritz John 条件) 设$\bar{\boldsymbol{x}}$为问题(3.4.1)的可行点，f和$g_i(i \in I(\bar{\boldsymbol{x}}))$在点$\bar{\boldsymbol{x}}$处可微，$h_j(j=1,2,\cdots,l)$在点$\bar{\boldsymbol{x}}$处具有一阶连续偏导数，$g_i(i \notin I(\bar{\boldsymbol{x}}))$在点$\bar{\boldsymbol{x}}$处连续，若$\bar{\boldsymbol{x}}$为问题(3.4.1)的局部极小点，则存在不全为0的数$u_0, u_i(i \in I(\bar{\boldsymbol{x}}))$和$v_j(j=1,2,\cdots,l)$，且$u_0 \geqslant 0, u_i \geqslant 0(i \in I(\bar{\boldsymbol{x}}))$，使

$$u_0 \nabla f(\bar{\boldsymbol{x}}) - \sum_{i \in I(\bar{\boldsymbol{x}})} u_i \nabla g_i(\bar{\boldsymbol{x}}) - \sum_{j=1}^{l} v_j \nabla h_j(\bar{\boldsymbol{x}}) = \boldsymbol{0}. \tag{3.4.5}$$

若$g_i(i \notin I(\bar{\boldsymbol{x}}))$在点$\bar{\boldsymbol{x}}$处也可微，则存在不全为0的数$u_0, u_i(i=1,2,\cdots,m)$和$v_j$ $(j=1,2,\cdots,l)$，且$u_0 \geqslant 0, u_i \geqslant 0(i=1,2,\cdots,m)$，使

$$\begin{cases} u_0 \nabla f(\bar{\boldsymbol{x}}) - \sum_{i=1}^{m} u_i \nabla g_i(\bar{\boldsymbol{x}}) - \sum_{j=1}^{l} v_j \nabla h_j(\bar{\boldsymbol{x}}) = \boldsymbol{0}, \\ u_i g_i(\bar{\boldsymbol{x}}) = 0, i = 1,2,\cdots,m. \end{cases} \tag{3.4.6}$$

证明 若$\nabla h_1(\bar{\boldsymbol{x}}), \nabla h_2(\bar{\boldsymbol{x}}), \cdots, \nabla h_l(\bar{\boldsymbol{x}})$线性相关，则存在不全为0的数$v_1, v_2, \cdots, v_l$，使

$$\sum_{j=1}^{l} v_j \nabla h_j(\bar{\boldsymbol{x}}) = \boldsymbol{0},$$

这时，可令$u_0=0, u_i=0(i \in I(\bar{\boldsymbol{x}}))$，得到(3.4.5)式. 下设$\nabla h_1(\bar{\boldsymbol{x}}), \nabla h_2(\bar{\boldsymbol{x}}), \cdots,$ $\nabla h_l(\bar{\boldsymbol{x}})$线性无关. 根据定理3.4.1，有

$$F_0 \cap G_0 \cap H_0 = \varnothing,$$

即关系式组

$$\begin{cases} \nabla f(\bar{\boldsymbol{x}})^{\mathrm{T}}\boldsymbol{d} < 0, \\ \nabla g_i(\bar{\boldsymbol{x}})^{\mathrm{T}}\boldsymbol{d} > 0, i \in I(\bar{\boldsymbol{x}}), \\ \nabla h_j(\bar{\boldsymbol{x}})^{\mathrm{T}}\boldsymbol{d} = 0, j = 1,2,\cdots,l \end{cases} \tag{3.4.7}$$

无解. 记$I(\bar{\boldsymbol{x}}) = \{i_1, i_2, \cdots, i_r\}$，且

$$\boldsymbol{A} = (\nabla f(\bar{\boldsymbol{x}}), -\nabla g_{i_1}(\bar{\boldsymbol{x}}), \cdots, -\nabla g_{i_r}(\bar{\boldsymbol{x}}))^{\mathrm{T}},$$

$$\boldsymbol{B} = (\nabla h_1(\bar{\boldsymbol{x}}), \nabla h_2(\bar{\boldsymbol{x}}), \cdots, \nabla h_l(\bar{\boldsymbol{x}}))^{\mathrm{T}},$$

则关系式组(3.4.7)无解等价于关系式组

$$\boldsymbol{A}\boldsymbol{d} < \boldsymbol{0},\ -\boldsymbol{B}\boldsymbol{d} = \boldsymbol{0} \tag{3.4.8}$$

无解.

由择一性定理 2.1.9 可知，关系式组(3.4.8)无解当且仅当存在 $\boldsymbol{u}\in\mathbb{R}^{r+1}$, $\boldsymbol{u}\geqslant\boldsymbol{0},\boldsymbol{u}\neq\boldsymbol{0}$ 及 $\boldsymbol{v}\in\mathbb{R}^{l}$,使得

$$\boldsymbol{A}^{\mathrm{T}}\boldsymbol{u}-\boldsymbol{B}^{\mathrm{T}}\boldsymbol{v}=\boldsymbol{0}. \tag{3.4.9}$$

从而,由 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 的意义以及(3.4.9)式知(3.4.5)式成立.

如果 $g_i(i\notin I(\bar{\boldsymbol{x}}))$ 在点 $\bar{\boldsymbol{x}}$ 处也可微,则只要令 $u_i=0(i\notin I(\bar{\boldsymbol{x}}))$,就得到(3.4.6)式. □

例 3.4.1 设

$$\begin{cases}\min & f(\boldsymbol{x})=x_1^2+x_2^2;\\ \text{s.t.} & g_1(\boldsymbol{x})=x_1^3-x_2\geqslant 0,\\ & g_2(\boldsymbol{x})=x_2\geqslant 0,\\ & h(\boldsymbol{x})=-(x_1-1)^2+x_2=0,\end{cases}$$

试判断点 $\bar{\boldsymbol{x}}=(1,0)^{\mathrm{T}}$ 是否为 Fritz John 点.

解 $I(\bar{\boldsymbol{x}})=\{2\}$,且

$$\nabla f(\bar{\boldsymbol{x}})=\begin{bmatrix}2\\0\end{bmatrix},\nabla g_2(\bar{\boldsymbol{x}})=\begin{bmatrix}0\\1\end{bmatrix},\nabla h(\bar{\boldsymbol{x}})=\begin{bmatrix}0\\1\end{bmatrix},$$

因此为使 Fritz John 条件

$$u_0\begin{bmatrix}2\\0\end{bmatrix}-u_2\begin{bmatrix}0\\1\end{bmatrix}-v\begin{bmatrix}0\\1\end{bmatrix}=\begin{bmatrix}0\\0\end{bmatrix}$$

成立,只有 $u_0=0$ 才行. 所以取 $u_0=0,u_2=1,v=-1$,即知 $\bar{\boldsymbol{x}}$ 是 Fritz John 点. □

图 3.4.1 中画出了这个最优化问题的可行域,它是抛物线 $x_2=(x_1-1)^2$ 上点 $\bar{\boldsymbol{x}}$ 右侧那部分线段.

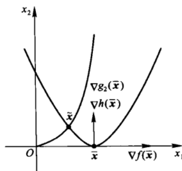


图 3.4.1 例 3.4.1 的图示

上例表明,在 Fritz John 条件中,不排除目标函数梯度的系数 $u_0=0$ 的情形.

为保证 $u_0 \neq 0$,需给约束条件施加某种约束规格,从而给出问题(3.4.1)的 K-T 必要条件.

定理 3.4.3(K-T 条件) 设 $\bar{x}$ 为问题(3.4.1)的可行点,f 和 $g_i(i \in I(\bar{x}))$ 在点 $\bar{x}$ 处可微,$h_j(j=1,2,\cdots,l)$ 在点 $\bar{x}$ 处具有一阶连续偏导数,$g_i(i \notin I(\bar{x}))$ 在点 $\bar{x}$ 处连续,且向量组

$$\nabla g_i(\bar{x})(i \in I(\bar{x})), \nabla h_j(\bar{x})(j = 1,2,\cdots,l)$$

线性无关. 若 $\bar{x}$ 是问题(3.4.1)的局部极小点,则存在数 $u_i \geqslant 0(i \in I(\bar{x}))$ 和 $v_j(j=1,2,\cdots,l)$,使

$$\nabla f(\bar{x}) - \sum_{i \in I(\bar{x})} u_i \nabla g_i(\bar{x}) - \sum_{j=1}^{l} v_j \nabla h_j(\bar{x}) = \mathbf{0}. \tag{3.4.10}$$

如果 $g_i(i \notin I(\bar{x}))$ 在点 $\bar{x}$ 处也可微,则存在数 $u_i \geqslant 0(i=1,2,\cdots,m)$ 和 $v_j(j=1,2,\cdots,l)$,使

$$\begin{cases} \nabla f(\bar{x}) - \sum_{i=1}^{m} u_i \nabla g_i(\bar{x}) - \sum_{j=1}^{l} v_j \nabla h_j(\bar{x}) = \mathbf{0}, \\ u_i g_i(\bar{x}) = 0, i = 1,2,\cdots,m. \end{cases} \tag{3.4.11}$$

证明 利用定理 3.4.2,仿照定理 3.3.4 的证明即可证明这个定理. □

K-T 条件(3.4.10)和(3.4.11)可看作是等式约束最优化问题(3.2.1)的 Lagrange 条件(3.2.2)与不等式约束最优化条件问题(3.3.1)的 K-T 条件的综合. 问题(3.4.1)比问题(3.3.1)多了 l 个等式约束,因此,相应的 K-T 条件多了 $-\sum_{j=1}^{l} v_j \nabla h_j(\bar{x})$ 这个项,而这恰好是 Lagrange 条件中等式约束对应的那一部分.

若记

$$\boldsymbol{g}(\boldsymbol{x}) = (g_1(\boldsymbol{x}), g_2(\boldsymbol{x}), \cdots, g_m(\boldsymbol{x}))^{\mathrm{T}},$$

$$\boldsymbol{h}(\boldsymbol{x}) = (h_1(\boldsymbol{x}), h_2(\boldsymbol{x}), \cdots, h_l(\boldsymbol{x}))^{\mathrm{T}},$$

$$\boldsymbol{u} = (u_1, u_2, \cdots, u_m)^{\mathrm{T}},$$

$$\boldsymbol{v} = (v_1, v_2, \cdots, v_l)^{\mathrm{T}},$$

则向量函数 $\boldsymbol{g}(\boldsymbol{x})$ 和 $\boldsymbol{h}(\boldsymbol{x})$ 的 Jacobi 矩阵分别为

$$\nabla \boldsymbol{g}(\boldsymbol{x}) = (\nabla g_1(\boldsymbol{x}), \nabla g_2(\boldsymbol{x}), \cdots, \nabla g_m(\boldsymbol{x}))^{\mathrm{T}},$$

$$\nabla \boldsymbol{h}(\boldsymbol{x}) = (\nabla h_1(\boldsymbol{x}), \nabla h_2(\boldsymbol{x}), \cdots, \nabla h_l(\boldsymbol{x}))^{\mathrm{T}},$$

从而 K-T 条件(3.4.11)可以写成

$$\begin{cases} \nabla f(\bar{\boldsymbol{x}}) - \nabla \boldsymbol{g}(\bar{\boldsymbol{x}})^{\mathrm{T}} \boldsymbol{u} - \nabla \boldsymbol{h}(\bar{\boldsymbol{x}})^{\mathrm{T}} \boldsymbol{v} = \mathbf{0}, \\ \boldsymbol{u}^{\mathrm{T}} \boldsymbol{g}(\bar{\boldsymbol{x}}) = 0, \\ \boldsymbol{u} \geqslant \mathbf{0}. \end{cases} \tag{3.4.12}$$

定义问题(3.4.1)的如下广义 Lagrange 函数(generalized Lagrange function)

$$L(\boldsymbol{x},\boldsymbol{u},\boldsymbol{v}) = f(\boldsymbol{x}) - \boldsymbol{u}^{\mathrm{T}}\boldsymbol{g}(\boldsymbol{x}) - \boldsymbol{v}^{\mathrm{T}}\boldsymbol{h}(\boldsymbol{x}),$$

$\boldsymbol{u},\boldsymbol{v}$ 称为广义 Lagrange 乘子向量(generalized Lagrange multipliers vector)或 K－T 乘子向量. 于是(3.4.12)式又可写作

$$\begin{cases}\nabla_x L(\bar{\boldsymbol{x}},\boldsymbol{u},\boldsymbol{v}) = \boldsymbol{0},\\ \boldsymbol{u}^{\mathrm{T}}\boldsymbol{g}(\bar{\boldsymbol{x}}) = 0,\\ \boldsymbol{u} \geqslant \boldsymbol{0}.\end{cases}$$

根据定理 3.4.3,若 $\bar{\boldsymbol{x}}$ 为问题(3.4.1)的局部极小点,则存在向量 $\bar{\boldsymbol{u}}\geqslant\boldsymbol{0}$ 和 $\bar{\boldsymbol{v}}$,使

$$\nabla_x L(\bar{\boldsymbol{x}},\bar{\boldsymbol{u}},\bar{\boldsymbol{v}}) = \boldsymbol{0},$$

即 $(\bar{\boldsymbol{x}}^{\mathrm{T}},\bar{\boldsymbol{u}}^{\mathrm{T}},\bar{\boldsymbol{v}}^{\mathrm{T}})^{\mathrm{T}}$ 为 $L(\boldsymbol{x},\boldsymbol{u},\boldsymbol{v})$ 的平稳点.

对于凸规划,上述 K－T 条件也是最优解的充分条件.

定理 3.4.4 设在问题(3.4.1)中,f 和 $-g_i(i=1,2,\cdots,m)$ 是凸函数,$h_j(j=1,2,\cdots,l)$ 是线性函数,$\bar{\boldsymbol{x}}$ 是可行点,并且 f 和 $g_i(i\in I(\bar{\boldsymbol{x}}))$ 在点 $\bar{\boldsymbol{x}}$ 处可微. 若 $\bar{\boldsymbol{x}}$ 是问题(3.4.1)的 K－T 点,则 $\bar{\boldsymbol{x}}$ 是问题(3.4.1)的全局极小点.

证明 由假设知问题(3.4.1)的可行域 S 是凸集. 因为 f 是凸函数且在点 $\bar{\boldsymbol{x}}$ 处可微,所以

$$f(\boldsymbol{x}) \geqslant f(\bar{\boldsymbol{x}}) + \nabla f(\bar{\boldsymbol{x}})^{\mathrm{T}}(\boldsymbol{x}-\bar{\boldsymbol{x}}),\ \forall \boldsymbol{x}\in S. \tag{3.4.13}$$

又 $g_i(i\in I(\bar{\boldsymbol{x}}))$ 为凹函数且在点 $\bar{\boldsymbol{x}}$ 处可微,故

$$g_i(\boldsymbol{x}) \leqslant g_i(\bar{\boldsymbol{x}}) + \nabla g_i(\bar{\boldsymbol{x}})^{\mathrm{T}}(\boldsymbol{x}-\bar{\boldsymbol{x}}),\ \forall i\in I(\bar{\boldsymbol{x}}),\ \forall \boldsymbol{x}\in S,$$

而 $\boldsymbol{x}\in S, i\in I(\bar{\boldsymbol{x}})$ 时,$g_i(\boldsymbol{x})\geqslant 0, g_i(\bar{\boldsymbol{x}})=0$. 于是

$$\nabla g_i(\bar{\boldsymbol{x}})^{\mathrm{T}}(\boldsymbol{x}-\bar{\boldsymbol{x}})\geqslant 0,\ \forall i\in I(\bar{\boldsymbol{x}}),\ \forall \boldsymbol{x}\in S. \tag{3.4.14}$$

由于 $h_j(\boldsymbol{x})(j=1,2,\cdots,l)$ 是线性函数,因此

$$h_j(\boldsymbol{x}) = h_j(\bar{\boldsymbol{x}}) + \nabla h_j(\bar{\boldsymbol{x}})^{\mathrm{T}}(\boldsymbol{x}-\bar{\boldsymbol{x}}),\ j=1,2,\cdots,l,\ \forall \boldsymbol{x}\in S.$$

因

$$h_j(\boldsymbol{x}) = h_j(\bar{\boldsymbol{x}}) = 0,\ j=1,2,\cdots,l,\ \forall \boldsymbol{x}\in S,$$

故

$$\nabla h_j(\bar{\boldsymbol{x}})^{\mathrm{T}}(\boldsymbol{x}-\bar{\boldsymbol{x}}) = 0,\ j=1,2,\cdots,l,\ \forall \boldsymbol{x}\in S. \tag{3.4.15}$$

注意到 $\bar{\boldsymbol{x}}$ 为 K－T 点,即存在数 $u_i\geqslant 0(i\in I(\bar{\boldsymbol{x}}))$ 和 $v_j(j=1,2,\cdots,l)$ 使

$$\nabla f(\bar{\boldsymbol{x}}) = \sum_{i\in I(\bar{\boldsymbol{x}})} u_i \nabla g_i(\bar{\boldsymbol{x}}) + \sum_{j=1}^{l} v_j \nabla h_j(\bar{\boldsymbol{x}}),$$

从而由(3.4.14)和(3.4.15)两式易知

$$\nabla f(\bar{\boldsymbol{x}})^{\mathrm{T}}(\boldsymbol{x}-\bar{\boldsymbol{x}}) \geqslant 0,\ \forall \boldsymbol{x}\in S.$$

把上式代入(3.4.13)式得

$$f(\boldsymbol{x}) \geqslant f(\bar{\boldsymbol{x}}),\ \forall \boldsymbol{x}\in S,$$

即 $\bar{\boldsymbol{x}}$ 为全局极小点. □

例 3.4.2 求解最优化问题

$$\begin{cases}\min & f(\boldsymbol{x}) = (x_1-3)^2+(x_2-1)^2,\\ \text{s.t.} & g(\boldsymbol{x}) = -x_1^2+x_2 \geqslant 0,\\ & h(\boldsymbol{x}) = 2x_1+x_2-3 = 0.\end{cases} \tag{3.4.16}$$

解 问题(3.4.16)的广义 Lagrange 函数为

$$\begin{aligned}L(\boldsymbol{x},u,v) &= f(\boldsymbol{x})-ug(\boldsymbol{x})-vh(\boldsymbol{x})\\ &= (x_1-3)^2+(x_2-1)^2-u(-x_1^2+x_2)-v(2x_1+x_2-3).\end{aligned}$$

因为

$$\frac{\partial L(\boldsymbol{x},u,v)}{\partial x_1} = 2(x_1-3)+2ux_1-2v,$$

$$\frac{\partial L(\boldsymbol{x},u,v)}{\partial x_2} = 2(x_2-1)-u-v,$$

所以问题(3.4.16)式的 K - T 条件及约束条件为

$$\begin{cases}2(x_1-3)+2ux_1-2v=0,\\ 2(x_2-1)-u-v=0,\\ u(-x_1^2+x_2)=0,\\ -x_1^2+x_2\geqslant 0,\\ 2x_1+x_2-3=0,\\ u\geqslant 0.\end{cases} \tag{3.4.17}$$

下面分两种情况讨论.

(1) 设 $u=0$,代入(3.4.17)的第 1、第 2 式连同第 5 式有

$$\begin{cases}2(x_1-3)-2v=0,\\ 2(x_2-1)-v=0,\\ 2x_1+x_2-3=0.\end{cases}$$

由此可解得 $x_1=\frac{7}{5}, x_2=\frac{1}{5}, v=-\frac{8}{5}$,但它们不满足(3.4.17)的第 4 式,即 $\boldsymbol{x}=\left(\frac{7}{5},\frac{1}{5}\right)^{\mathrm{T}}$ 不是问题(3.4.16)式的可行点,因而不是 K - T 点.

(2) 设 $u>0$,此时(3.4.17)式成为

$$\begin{cases}2(x_1-3)+2ux_1-2v=0,\\ 2(x_2-1)-u-v=0,\\ -x_1^2+x_2=0,\\ 2x_1+x_2-3=0.\end{cases} \tag{3.4.18}$$

由(3.4.18)的第 3 式、第 4 式有

$$-x_1^2-2x_1+3=0,$$

解得 $x_1=1$ 或 $x_1=-3$，代入(3.4.18)的第 4 式，得 $x_2=1$ 或 $x_2=9$，再分别代入(3.4.18)的第 1 式和第 2 式，得 $u=1$ 或 $u=-11$（此与 $u>0$ 矛盾），$v=-1$ 或 $v=27$. 由此可知 $\bar{\boldsymbol{x}}=(1,1)^{\mathrm{T}}$ 是问题(3.4.16)的 K－T 点，但 $\boldsymbol{x}=(-3,9)^{\mathrm{T}}$ 不是问题的 K－T 点.

容易验证，f 是 $\mathbb{R}^2$ 上的凸函数，g 是 $\mathbb{R}^2$ 上的凹函数，h 是线性函数，因此由定理 3.4.4 知，$\bar{\boldsymbol{x}}=(1,1)^{\mathrm{T}}$ 是问题(3.4.16)的全局最优解（见图 3.4.2）. □

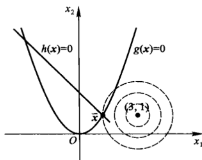


图 3.4.2　例 3.4.2 的图示

利用广义 Lagrange 函数并仿照定理 3.2.2 可以给出问题(3.4.1)的局部最优解的二阶充分条件.

定理 3.4.5　设 $\bar{\boldsymbol{x}}$ 为问题(3.4.1)的可行点，$f, g_i\,(i\in I(\bar{\boldsymbol{x}}))$ 和 $h_j\,(j=1,2,\cdots,l)$ 在点 $\bar{\boldsymbol{x}}$ 处具有二阶连续偏导数，并且存在乘子向量 $\bar{\boldsymbol{u}}=(\bar{u}_1,\bar{u}_2,\cdots,\bar{u}_m)^{\mathrm{T}}\geqslant\boldsymbol{0}$ 和 $\bar{\boldsymbol{v}}=(\bar{v}_1,\bar{v}_2,\cdots,\bar{v}_m)^{\mathrm{T}}$ 使 K－T 条件成立，即

$$\begin{cases}\nabla_x L(\bar{\boldsymbol{x}},\bar{\boldsymbol{u}},\bar{\boldsymbol{v}})=\boldsymbol{0},\\ \bar{\boldsymbol{u}}^{\mathrm{T}}\boldsymbol{g}(\bar{\boldsymbol{x}})=0.\end{cases}\tag{3.4.19}$$

若对于任何满足

$$\begin{cases}\boldsymbol{z}^{\mathrm{T}}\nabla g_i(\bar{\boldsymbol{x}})\geqslant 0, i\in I(\bar{\boldsymbol{x}})\text{ 且 }\bar{u}_i=0,\\ \boldsymbol{z}^{\mathrm{T}}\nabla g_i(\bar{\boldsymbol{x}})=0, i\in I(\bar{\boldsymbol{x}})\text{ 且 }\bar{u}_i>0,\\ \boldsymbol{z}^{\mathrm{T}}\nabla h_j(\bar{\boldsymbol{x}})=0, j=1,2,\cdots,l\end{cases}\tag{3.4.20}$$

的向量 $\boldsymbol{z}\neq\boldsymbol{0}$，都有

$$\boldsymbol{z}^{\mathrm{T}}\nabla_{xx}^2 L(\bar{\boldsymbol{x}},\bar{\boldsymbol{u}},\bar{\boldsymbol{v}})\boldsymbol{z}>0,\tag{3.4.21}$$

则 $\bar{\boldsymbol{x}}$ 为问题(3.4.1)的严格局部极小点.

证明　用反证法. 假设 $\bar{\boldsymbol{x}}$ 不是问题(3.4.1)的严格局部极小点，则在 $\bar{\boldsymbol{x}}$ 的某邻域内存在收敛于 $\bar{\boldsymbol{x}}$ 的点列 $\{\boldsymbol{x}_k\}\subset S$，$\boldsymbol{x}_k\neq\bar{\boldsymbol{x}}$，$k=1,2,\cdots$，使

$$f(\boldsymbol{x}_k)\leqslant f(\bar{\boldsymbol{x}}), k=1,2,\cdots.\tag{3.4.22}$$

令

$$z_k = \frac{\boldsymbol{x}_k - \bar{\boldsymbol{x}}}{\|\boldsymbol{x}_k - \bar{\boldsymbol{x}}\|}, k = 1,2,\cdots,$$

则$\{z_k\}$有界,因此有收敛子列,不妨设$\{z_k\}$收敛,其极限记为z,从而$\|z\| = 1$.根据 Taylor 公式,有

$$g_i(\boldsymbol{x}_k) = g_i(\bar{\boldsymbol{x}}) + \nabla g_i(\bar{\boldsymbol{x}})^{\mathrm{T}}(\boldsymbol{x}_k - \bar{\boldsymbol{x}}) + o(\|\boldsymbol{x}_k - \bar{\boldsymbol{x}}\|). \tag{3.4.23}$$

当$i \in I(\bar{\boldsymbol{x}})$时,$g_i(\bar{\boldsymbol{x}}) = 0$;而$\boldsymbol{x}_k \in S$,故$g_i(\boldsymbol{x}_k) \geqslant 0$. 于是由(3.4.23)式得

$$\nabla g_i(\bar{\boldsymbol{x}})^{\mathrm{T}}(\boldsymbol{x}_k - \bar{\boldsymbol{x}}) + o(\|\boldsymbol{x}_k - \bar{\boldsymbol{x}}\|) \geqslant 0. \tag{3.4.23$'$}$$

上式两边除以$\|\boldsymbol{x}_k - \bar{\boldsymbol{x}}\|$,并令$k \to \infty$,得

$$\boldsymbol{z}^{\mathrm{T}} \nabla g_i(\bar{\boldsymbol{x}}) \geqslant 0, \forall i \in I(\bar{\boldsymbol{x}}). \tag{3.4.24}$$

用类似方法不难得到

$$\boldsymbol{z}^{\mathrm{T}} \nabla h_j(\bar{\boldsymbol{x}}) = 0, j = 1,2,\cdots,l, \tag{3.4.25}$$

和

$$\boldsymbol{z}^{\mathrm{T}} \nabla f(\bar{\boldsymbol{x}}) \leqslant 0. \tag{3.4.26}$$

下面分两种情况讨论:

(1) $\boldsymbol{z}$不满足(3.4.20)式

此时,由式(3.4.24)和(3.4.20)可知,存在$i \in I(\bar{\boldsymbol{x}})$,使

$$\bar{u}_i > 0, \boldsymbol{z}^{\mathrm{T}} \nabla g_i(\bar{\boldsymbol{x}}) > 0,$$

从而由 K-T 条件(3.4.19)及(3.4.24)、(3.4.25)式和$\bar{\boldsymbol{u}} \geqslant \boldsymbol{0}$,得

$$\boldsymbol{z}^{\mathrm{T}} \nabla f(\bar{\boldsymbol{x}}) = \boldsymbol{z}^{\mathrm{T}} \left(\sum_{i \in I(\bar{\boldsymbol{x}})} \bar{u}_i \nabla g_i(\bar{\boldsymbol{x}}) + \sum_{j=1}^{l} \bar{v}_j \nabla h_j(\bar{\boldsymbol{x}}) \right) > 0,$$

此与(3.4.26)式相矛盾.

(2) $\boldsymbol{z}$满足(3.4.20)式

这时,把广义 Lagrange 函数$L(\boldsymbol{x},\bar{\boldsymbol{u}},\bar{\boldsymbol{v}})$在点$\bar{\boldsymbol{x}}$处应用二阶 Taylor 公式,有

$$L(\boldsymbol{x}_k,\bar{\boldsymbol{u}},\bar{\boldsymbol{v}}) = L(\bar{\boldsymbol{x}},\bar{\boldsymbol{u}},\bar{\boldsymbol{v}}) + \nabla_x L(\bar{\boldsymbol{x}},\bar{\boldsymbol{u}},\bar{\boldsymbol{v}})^{\mathrm{T}}(\boldsymbol{x}_k - \bar{\boldsymbol{x}}) + \frac{1}{2}(\boldsymbol{x}_k - \bar{\boldsymbol{x}})^{\mathrm{T}} \nabla_{xx}^2 L(\bar{\boldsymbol{x}},\bar{\boldsymbol{u}},\bar{\boldsymbol{v}})(\boldsymbol{x}_k - \bar{\boldsymbol{x}}) + o(\|\boldsymbol{x}_k - \bar{\boldsymbol{x}}\|^2), \tag{3.4.27}$$

因为$\boldsymbol{x}_k \in S$,$\bar{\boldsymbol{u}} \geqslant \boldsymbol{0}$,且由广义 Lagrange 函数$L(\boldsymbol{x},\boldsymbol{u},\boldsymbol{v})$的定义知,对每个$\boldsymbol{x}_k$,有

$$L(\boldsymbol{x}_k,\bar{\boldsymbol{u}},\bar{\boldsymbol{v}}) = f(\boldsymbol{x}_k) - \sum_{i=1}^{m} \bar{u}_i g_i(\boldsymbol{x}_k) - \sum_{j=1}^{l} \bar{v}_j h_j(\boldsymbol{x}_k),$$

所以

$$L(\boldsymbol{x}_k,\bar{\boldsymbol{u}},\bar{\boldsymbol{v}}) \leqslant f(\boldsymbol{x}_k), k = 1,2,\cdots, \tag{3.4.28}$$

且

$$L(\bar{\boldsymbol{x}},\bar{\boldsymbol{u}},\bar{\boldsymbol{v}}) = f(\bar{\boldsymbol{x}}). \tag{3.4.29}$$

又由(3.4.19)式有

$$\nabla_x L(\bar{\boldsymbol{x}},\bar{\boldsymbol{u}},\bar{\boldsymbol{v}}) = \boldsymbol{0}, \tag{3.4.30}$$

于是把式(3.4.28)~(3.4.30)和式(3.4.22)代入(3.4.27)式,得

$$\frac{1}{2}(\boldsymbol{x}_k-\bar{\boldsymbol{x}})^{\mathrm{T}}\nabla_{xx}^2L(\bar{\boldsymbol{x}},\bar{\boldsymbol{u}},\bar{\boldsymbol{v}})(\boldsymbol{x}_k-\bar{\boldsymbol{x}})+o(\|\boldsymbol{x}_k-\bar{\boldsymbol{x}}\|^2)\leqslant 0.$$

上式两边除以 $\|\boldsymbol{x}_k-\bar{\boldsymbol{x}}\|^2$,并令 $k\to\infty$,得到

$$\boldsymbol{z}^{\mathrm{T}}\nabla_{xx}^2L(\bar{\boldsymbol{x}},\bar{\boldsymbol{u}},\bar{\boldsymbol{v}})\boldsymbol{z}\leqslant 0,$$

这与(3.4.21)式相矛盾. □

例 3.4.3 求解下面最优化问题

$$\begin{cases}\min & f(\boldsymbol{x})=\sum_{i=1}^{n}\dfrac{c_i}{x_i};\\ \text{s.t.} & g_i(\boldsymbol{x})=x_i\geqslant 0, i=1,2,\cdots,n,\\ & h(\boldsymbol{x})=\sum_{i=1}^{n}a_ix_i-b=0,\end{cases}\tag{3.4.31}$$

其中常数 $a_i>0, c_i>0, i=1,2,\cdots,n; b>0$.

解 问题(3.4.31)的广义 Lagrange 函数为

$$L(\boldsymbol{x},\boldsymbol{u},v)=\sum_{i=1}^{n}\frac{c_i}{x_i}-\sum_{i=1}^{n}u_ix_i-v\left(\sum_{i=1}^{n}a_ix_i-b\right),$$

因为

$$\frac{\partial L(\boldsymbol{x},\boldsymbol{u},v)}{\partial x_i}=-\frac{c_i}{x_i^2}-u_i-a_iv, i=1,2,\cdots,n,$$

所以问题(3.4.31)的 K-T 条件及约束条件为

$$\begin{cases}-\dfrac{c_i}{x_i^2}-u_i-a_iv=0, i=1,2,\cdots,n,\\ u_ix_i=0, i=1,2,\cdots,n,\\ x_i\geqslant 0, i=1,2,\cdots,n,\\ \sum_{i=1}^{n}a_ix_i-b=0,\\ u_i\geqslant 0, i=1,2,\cdots,n.\end{cases}\tag{3.4.32}$$

由(3.4.32)的第 1 式、第 3 式知 $x_i>0(i=1,2,\cdots,n)$,从而由(3.4.32)的第 2 式解得

$$\bar{u}_i=0, i=1,2,\cdots,n.$$

于是由(3.4.32)的第 1 式知 $v<0$,且

$$a_ivx_i^2+c_i=0, i=1,2,\cdots,n,$$

即得

$$x_i=\sqrt{\frac{c_i}{-a_iv}}, i=1,2,\cdots,n.\tag{3.4.33}$$

将上式代入(3.4.32)的第4式,得

$$\sum_{i=1}^{n} a_i \sqrt{\frac{c_i}{-a_i v}} - b = 0,$$

解得

$$\bar{v} = -\frac{\left(\sum_{i=1}^{n} \sqrt{a_i c_i}\right)^2}{b^2},$$

代入(3.4.33)式即有

$$\bar{x}_i = \frac{b}{\sum_{i=1}^{n} \sqrt{a_i c_i}} \sqrt{\frac{c_i}{a_i}}, i = 1,2,\cdots,n.$$

所以 $\bar{\boldsymbol{x}} = (\bar{x}_1, \bar{x}_2, \cdots, \bar{x}_n)^{\mathrm{T}}$ 是问题(3.4.31)的 K-T 点.

又由于 $L(\boldsymbol{x}, \bar{\boldsymbol{u}}, \bar{v})$ 在点 $(\bar{\boldsymbol{x}}^{\mathrm{T}}, \bar{\boldsymbol{u}}^{\mathrm{T}}, \bar{v})^{\mathrm{T}}$ 处关于 $\boldsymbol{x}$ 的 Hesse 矩阵 $\nabla_{xx}^2 L(\bar{\boldsymbol{x}}, \bar{\boldsymbol{u}}, \bar{v})$ 是一个 n 阶对角矩阵,其对角线上第 i 个元素为

$$\frac{2c_i}{\bar{x}_i^3} > 0, i = 1,2,\cdots,n,$$

因此 $\nabla_{xx}^2 L(\bar{\boldsymbol{x}}, \bar{\boldsymbol{u}}, \bar{v})$ 是正定矩阵. 根据定理 3.4.5, $\bar{\boldsymbol{x}}$ 为问题(3.4.31)的严格局部极小点. □

习　题　三

1. 试讨论参数 α 取何值时, $\mathbb{R}^3$ 的原点是无约束最优化问题

$$\min \ \alpha x_1^2 e^{x_2} + x_2^2 e^{x_3} + x_3^2 e^{x_1}$$

的最优解?

2. 把下列二次函数 $f(\boldsymbol{x})$ 写成

$$f(\boldsymbol{x}) = \frac{1}{2}\boldsymbol{x}^{\mathrm{T}} \boldsymbol{Q} \boldsymbol{x} + \boldsymbol{b}^{\mathrm{T}} \boldsymbol{x} + c$$

的形式,并求无约束最优化问题 $\min\limits_{\boldsymbol{x} \in \mathbb{R}^n} f(\boldsymbol{x})$ 的最优解和最优值.

(1) $f(x_1, x_2) = x_1^2 + 2x_2^2 + x_1 x_2$;

(2) $f(x_1, x_2, x_3) = 3x_1^2 + 4x_2^2 + x_3^2 - 2x_1 x_2 + 3x_2 x_3 - 4x_1 + 5x_3 + 10$.

3. 试求解下列等式约束最优化问题:

(1) $\begin{cases} \min & x_1^2 + 4x_2^2 + 16x_3^2; \\ \text{s.t.} & x_1 x_2 = 1. \end{cases}$

(2) $\begin{cases} \min & x_1 + x_2 + x_3; \\ \text{s.t.} & \dfrac{a}{x_1} + \dfrac{b}{x_2} + \dfrac{c}{x_3} = 1 (a, b, c > 0). \end{cases}$

(3) $\begin{cases}\min & \sum_{i=1}^{n} x_i^p;\\ \text{s.t.} & \sum_{i=1}^{n} x_i=\alpha(p>1,\alpha>0).\end{cases}$

4. 证明等式约束最优化问题

$$\begin{cases}\min & x_1^2+x_2^2;\\ \text{s.t.} & (x_1-1)^3-x_2^2=0\end{cases}$$

不存在局部最优解.

5. 验证点 $\bar{x}=(3,1)^{\mathrm{T}}$ 是不等式约束最优化问题

$$\begin{cases}\min & (x_1-7)^2+(x_2-3)^2;\\ \text{s.t.} & 10-x_1^2-x_2^2\geqslant 0,\\ & 4-x_1-x_2\geqslant 0,\\ & x_2\geqslant 0\end{cases}$$

的 Fritz John 点.

6. 求下列约束最优化问题的 K－T 点:

(1) $\begin{cases}\min & \dfrac{1}{x_1x_2x_3};\\ \text{s.t.} & 72-x_1-2x_2-2x_3\geqslant 0,\\ & x_1>0,x_2>0,x_3>0.\end{cases}$

(2) $\begin{cases}\min & -3x_1+x_2-x_3^2;\\ \text{s.t.} & x_1+x_2+x_3\leqslant 0,\\ & -x_1+2x_2+x_3^2=0.\end{cases}$

7. 设 $f:\mathbb{R}^n\to\mathbb{R}$ 为可微函数,证明问题

$$\begin{cases}\min & f(\boldsymbol{x});\\ \text{s.t.} & \boldsymbol{x}\geqslant\boldsymbol{0}\end{cases}$$

的 K－T 条件是

$$\begin{cases}\nabla f(\bar{\boldsymbol{x}})^{\mathrm{T}}\bar{\boldsymbol{x}}=0,\\ \nabla f(\bar{\boldsymbol{x}})\geqslant\boldsymbol{0}.\end{cases}$$

并说明这个条件的几何意义.

8. 设 $\boldsymbol{c}\in\mathbb{R}^n$ 为非零向量,求解不等式约束最优化问题

$$\begin{cases}\max & \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x};\\ \text{s.t.} & \boldsymbol{x}^{\mathrm{T}}\boldsymbol{x}\leqslant 1.\end{cases}$$

9. 设 $\alpha>0$,求解约束最优化问题

$$\begin{cases}\min & \sum_{j=1}^{n} x_j^2;\\ \text{s.t.} & \prod_{j=1}^{n} x_j=\alpha,\\ & x_j\geqslant 0,j=1,2,\cdots,n.\end{cases}$$

10. 设 $\boldsymbol{Q} \in \mathbb{R}^{n \times n}$ 为正定矩阵，$\boldsymbol{A} \in \mathbb{R}^{m \times n}$ 为行满秩矩阵，试求问题

$$\begin{cases} \max & \frac{1}{2}\boldsymbol{x}^{\mathrm{T}}\boldsymbol{Q}\boldsymbol{x}; \\ \text{s.t.} & \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}, \\ & \boldsymbol{x} \geqslant \boldsymbol{0} \end{cases}$$

的 K－T 点，并问它是否为全局最优解？

第四章　线性规划

线性规划在数学规划中占有重要的地位,这不仅因为它相对地简单,在理论和算法上都比较成熟,还因为它本身在实际问题中有着广泛的应用,并且能为求解某些非线性规划问题所引证.

本章介绍线性规划的基本理论、单纯形法、对偶理论和对偶单纯形法.

4.1　线性规划的基本理论

线性规划问题就是一个线性函数在线性等式或不等式约束条件下的极值问题,其数学模型的一般形式(general form)为

$$\begin{cases} \min/\max \sum_{j=1}^{n} c_j x_j; \\ \text{s.t.} \quad \sum_{j=1}^{n} a_{ij} x_j \leqslant b_i (\text{或} \geqslant b_i, \text{或} = b_i), i = 1,2,\cdots,m. \end{cases}$$

为了便于讨论线性规划问题的一般解法,常将线性规划问题用统一的形式表示为

$$\begin{cases} \min \quad \sum_{j=1}^{n} c_j x_j; \\ \text{s.t.} \quad \sum_{j=1}^{n} a_{ij} x_j = b_i, i = 1,2,\cdots,m, \\ \qquad x_j \geqslant 0, j = 1,2,\cdots,n, \end{cases} \tag{4.1.1}$$

式(4.1.1)称为线性规划模型的标准形式(standard form),并简记为(LP).不难知道,这种标准形式是不失一般性的,因为任何一个线性规划问题都能化为标准形式.

利用向量和矩阵的记号,可以把(LP)写成如下的形式

$$\begin{cases} \min \quad \boldsymbol{c}^{\mathrm{T}} \boldsymbol{x}; \\ \text{s.t.} \quad \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}, \\ \qquad \boldsymbol{x} \geqslant \boldsymbol{0}, \end{cases} \tag{LP}$$

其中

$$x=\begin{bmatrix}x_1\\x_2\\\vdots\\x_n\end{bmatrix},c=\begin{bmatrix}c_1\\c_2\\\vdots\\c_n\end{bmatrix},b=\begin{bmatrix}b_1\\b_2\\\vdots\\b_m\end{bmatrix},A=\begin{bmatrix}a_{11}&a_{12}&\cdots&a_{1n}\\a_{21}&a_{22}&\cdots&a_{2n}\\\vdots&\vdots&&\vdots\\a_{m1}&a_{m2}&\cdots&a_{mn}\end{bmatrix},$$

矩阵 $\boldsymbol{A}$ 称为约束矩阵(constraint matrix);$\boldsymbol{Ax}=\boldsymbol{b}$ 称为约束方程组(system of constraint equations);$\boldsymbol{x}\geqslant\boldsymbol{0}$ 称为非负约束(nonnegative constraint).

为了方便,除特别说明外,我们总是假定:在(LP)中,$n\geqslant m\geqslant 1$;$\boldsymbol{A}$ 是行满秩的,即 rank $\boldsymbol{A}=m$.

定义 4.1.1 在(LP)中,满足约束方程组及非负约束的向量 $\boldsymbol{x}$ 称为可行解(feasible solution)或可行点(feasible point);所有可行解的全体称为可行解集(feasible solutions set)或可行域(feasible region),记作 K,即

$$K=\{\boldsymbol{x}\mid \boldsymbol{Ax}=\boldsymbol{b},\boldsymbol{x}\geqslant\boldsymbol{0}\};$$

使目标函数在 K 上取到最小值的可行解称为最优解(optimal solution);最优解对应的目标函数值称为最优值(optimal value).

在第二章 2.2 节中,我们从几何理论和直观上考察了(LP)的可行域 K,给出了 K 的表示定理. 为了求解(LP)的实际需要,还必须引进一些有关的代数概念.

定义 4.1.2 在(LP)中,约束矩阵 $\boldsymbol{A}$ 的任意一个 m 阶满秩子方阵 $\boldsymbol{B}$ 称为基(basis);$\boldsymbol{B}$ 中 m 个线性无关的列向量称为基向量(basic vectors),$\boldsymbol{x}$ 中与 $\boldsymbol{B}$ 的列对应的分量称为关于 $\boldsymbol{B}$ 的基变量(basic variables),其余的变量称为关于 $\boldsymbol{B}$ 的非基变量(nonbasic variables).

由于 rank $\boldsymbol{A}=m$,因此,在(LP)中总存在基,并且基的个数不超过$\binom{n}{m}$.

任取(LP)的一个基

$$\boldsymbol{B}=(\boldsymbol{p}_{j_1},\boldsymbol{p}_{j_2},\cdots,\boldsymbol{p}_{j_m}),$$

此时关于基 $\boldsymbol{B}$ 的所有基变量构成的向量记为

$$\boldsymbol{x}_B=(x_{j_1},x_{j_2},\cdots,x_{j_m})^{\mathrm{T}},$$

若令关于 $\boldsymbol{B}$ 的非基变量都取 0,则约束方程 $\boldsymbol{Ax}=\boldsymbol{b}$ 变为

$$\boldsymbol{Bx}_B=\boldsymbol{b}. \tag{4.1.2}$$

由于 $\boldsymbol{B}$ 是满秩方阵,因此(4.1.2)式有惟一的解

$$\boldsymbol{x}_B=\boldsymbol{B}^{-1}\boldsymbol{b}.$$

记 $\boldsymbol{B}^{-1}\boldsymbol{b}=(\bar{x}_{j_1},\bar{x}_{j_2},\cdots,\bar{x}_{j_m})^{\mathrm{T}}$,则由

$$x_{j_k}=\bar{x}_{j_k},k=1,2,\cdots,m,$$
$$x_j=0,\forall j\in\{1,2,\cdots,n\}\setminus\{j_1,j_2,\cdots,j_m\}$$

所构成的 n 维向量 $\bar{x}$ 是 $Ax=b$ 的一个解，称之为(LP)的关于 B 的基本解(basic solution).

基本解满足约束方程组，但不一定满足非负约束，所以不一定是可行解. 若 $B^{-1}b\geqslant 0$，即基本解 $\bar{x}$ 也是可行解，则称 $\bar{x}$ 为(LP)的关于基 B 的基本可行解(basic feasible soloution)，相应的基 B 称为(LP)的可行基(feasible basis)；当 $B^{-1}b>0$ 时，称此基本可行解 $\bar{x}$ 是非退化的(nondegenerate)，否则，称之为退化的(degenerate). 若一个(LP)的所有基本可行解都是非退化的，则称该(LP)是非退化的，否则，称它是退化的.

例 4.1.1 求下面线性规划问题的所有基本可行解

$$\begin{cases}\min & 4x_1-4x_2;\\ \text{s.t.} & x_1-x_2+x_3=4,\\ & -x_1+x_2+x_4=2,\\ & x_j\geqslant 0, j=1,2,3,4.\end{cases}\tag{4.1.3}$$

解 问题(4.1.3)的约束矩阵的4个列向量依次为

$$p_1=\begin{bmatrix}1\\-1\end{bmatrix}, p_2=\begin{bmatrix}-1\\1\end{bmatrix}, p_3=\begin{bmatrix}1\\0\end{bmatrix}, p_4=\begin{bmatrix}0\\1\end{bmatrix},$$

不难得到问题(4.1.3)的全部基

$$B_1=(p_1,p_3), B_2=(p_1,p_4), B_3=(p_2,p_3),$$
$$B_4=(p_2,p_4), B_5=(p_3,p_4).$$

对于 B_1，x_1 和 x_3 为基变量，x_2 和 x_4 为非基变量. 令 $x_2=x_4=0$，有

$$\begin{cases}x_1+x_3=4,\\ -x_1=2,\end{cases}$$

得到关于 B_1 的基本解 $x_1=(-2,0,6,0)^{\mathrm{T}}$，它不是可行解.

对于 B_2，x_1 和 x_4 为基变量，x_2 和 x_3 为非基变量. 令 $x_2=x_3=0$，有

$$\begin{cases}x_1=4,\\ -x_1+x_4=2,\end{cases}$$

得到关于 B_2 的基本解 $x_2=(4,0,0,6)^{\mathrm{T}}$，它是一个非退化的基本可行解.

同理，可求得关于 B_3,B_4,B_5 的基本解分别为

$$x_3=(0,2,6,0)^{\mathrm{T}}, x_4=(0,-4,0,6)^{\mathrm{T}}, x_5=(0,0,4,2)^{\mathrm{T}}.$$

显然，x_3 和 x_5 均是非退化的基本可行解，而 x_4 不是可行解. 因此，问题(4.1.3)的所有基本可行解为 x_2,x_3,x_5. 此外，因为问题(4.1.3)的所有基本可行解都是非退化的，所以问题(4.1.3)是非退化的. □

下面的定理给出了判别一个可行解是否为基本可行解的一个准则.

定理 4.1.1 设 $\bar{x}$ 为(LP)的可行解，则 $\bar{x}$ 为(LP)的基本可行解的充要条件

是它的非零分量所对应的列向量线性无关.

证明 不妨设 $\bar{x}$ 的前 r 个分量为正分量,即

$$\bar{x} = (\bar{x}_1,\bar{x}_2,\cdots,\bar{x}_r,0,\cdots,0)^{\mathrm{T}},\bar{x}_j > 0(j = 1,2,\cdots,r).$$

若 $\bar{x}$ 是基本可行解,则取正值的变量 $x_1,x_2,\cdots,x_r$ 必定是基变量,而这些基变量对应的列向量 $\boldsymbol{p}_1,\boldsymbol{p}_2,\cdots,\boldsymbol{p}_r$ 是基向量,故必定线性无关.

反之,若 $\boldsymbol{p}_1,\boldsymbol{p}_2,\cdots,\boldsymbol{p}_r$ 线性无关,则必有 $0\leqslant r\leqslant m$,当 $r = m$ 时,$\boldsymbol{B} = (\boldsymbol{p}_1,\boldsymbol{p}_2,\cdots,\boldsymbol{p}_r)$就是一个基;当 $r<m$ 时,一定可以从约束矩阵 $\boldsymbol{A}$ 的后 $n-r$ 个列向量中选出 $m-r$ 个,不妨设为 $\boldsymbol{p}_{r+1},\boldsymbol{p}_{r+2},\cdots,\boldsymbol{p}_m$,使

$$\boldsymbol{B} = (\boldsymbol{p}_1,\cdots,\boldsymbol{p}_r,\boldsymbol{p}_{r+1},\cdots,\boldsymbol{p}_m)$$

成为一个基,由于 $\bar{x}$ 是可行解,因此

$$\sum_{j=1}^{r}\bar{x}_j\boldsymbol{p}_j = \boldsymbol{b},$$

从而必有

$$\sum_{j=1}^{m}\bar{x}_j\boldsymbol{p}_j = \boldsymbol{b}.$$

由此可知 $\bar{x}$ 是关于 $\boldsymbol{B}$ 的基本可行解. □

由定理 4.1.1 和定理 2.2.2,立即得到

定理 4.1.2 $\bar{x}$ 是(LP)的基本可行解的充要条件是 $\bar{x}$ 为(LP)的可行域的极点. □

需要指出的是,一个基本可行解对应着惟一的一个极点,而一个极点可能对应着几个不同的基本可行解. 这就是说,(LP)的基本可行解与可行域 K 的极点不是一一对应的.

例 4.1.2 求下面线性规划问题的可行域的极点

$$\begin{cases}\min & x_1 - x_2;\\ \text{s.t.} & x_1 + 2x_2 + x_3 = 2,\\ & x_1 + x_4 = 2,\\ & x_j \geqslant 0, j = 1,2,3,4.\end{cases}\tag{4.1.4}$$

解 因为问题(4.1.4)的约束矩阵的 4 个列向量依次为

$$\boldsymbol{p}_1 = \begin{bmatrix}1\\1\end{bmatrix},\boldsymbol{p}_2 = \begin{bmatrix}2\\0\end{bmatrix},\boldsymbol{p}_3 = \begin{bmatrix}1\\0\end{bmatrix},\boldsymbol{p}_4 = \begin{bmatrix}0\\1\end{bmatrix},$$

不难得知问题(4.1.4)的基的全部为

$$\boldsymbol{B}_1 = (\boldsymbol{p}_1,\boldsymbol{p}_2),\boldsymbol{B}_2 = (\boldsymbol{p}_1,\boldsymbol{p}_3),\boldsymbol{B}_3 = (\boldsymbol{p}_1,\boldsymbol{p}_4),$$
$$\boldsymbol{B}_4 = (\boldsymbol{p}_2,\boldsymbol{p}_4),\boldsymbol{B}_5 = (\boldsymbol{p}_3,\boldsymbol{p}_4).$$

所以容易求得关于基 $\boldsymbol{B}_1,\boldsymbol{B}_2,\boldsymbol{B}_3,\boldsymbol{B}_4,\boldsymbol{B}_5$ 的基本解分别为

$$\boldsymbol{x}_1 = (2,0,0,0)^{\mathrm{T}},\boldsymbol{x}_2 = (2,0,0,0)^{\mathrm{T}},\boldsymbol{x}_3 = (2,0,0,0)^{\mathrm{T}},$$

$$x_4 = (0,1,0,2)^{\mathrm{T}}, x_5 = (0,0,2,2)^{\mathrm{T}}.$$

显然,x_1,x_2,x_3 均为退化的基本可行解,x_4,x_5 是非退化的基本可行解. 因此问题(4.1.4)的可行域有三个极点:$(2,0,0,0)^{\mathrm{T}},(0,1,0,2)^{\mathrm{T}},(0,0,2,2)^{\mathrm{T}}$.

由上可知,问题(4.1.4)是退化的,并且关于三个不同的基 $\boldsymbol{B}_1,\boldsymbol{B}_2,\boldsymbol{B}_3$ 的基本可行解 x_1,x_2,x_3 对应着同一个极点$(2,0,0,0)^{\mathrm{T}}$. □

根据定理 2.2.1 和定理 4.1.2,显然有

定理 4.1.3 若(LP)有可行解,则它必有基本可行解. □

定理 4.1.4 若(LP)可行域 K 非空有界,则(LP)必存在最优解,且其中至少有一个基本可行解为最优解.

证明 根据推论 2.2.6,(LP)的任一可行解 x 都可表示为(LP)的全部基本可行解 $x_1,x_2,\cdots,x_k$ 的凸组合,即

$$x = \sum_{i=1}^{k} \lambda_i x_i, \ \forall x \in K,$$

其中 $\lambda_i \geqslant 0(i=1,2,\cdots,k)$, $\sum_{i=1}^{k} \lambda_i = 1$.

设 x_s 是使(LP)中目标函数值达到最小的基本可行解,即

$$c^{\mathrm{T}} x_s = \min_{1 \leqslant i \leqslant k} c^{\mathrm{T}} x_i,$$

则

$$c^{\mathrm{T}} x = \sum_{i=1}^{k} \lambda_i c^{\mathrm{T}} x_i \geqslant \sum_{i=1}^{k} \lambda_i c^{\mathrm{T}} x_s = c^{\mathrm{T}} x_s, \ \forall x \in K.$$

这表明,基本可行解 x_s 为(LP)的最优解. □

定理 4.1.5 设(LP)的可行域 K 无界,则(LP)存在最优解的充要条件是对 K 的任一极方向 d,均有 $c^{\mathrm{T}} d \geqslant 0$.

证明 根据定理 2.2.10,(LP)的任一可行解 x 都可写成

$$x = \sum_{i=1}^{k} \lambda_i x_i + \sum_{j=1}^{l} \mu_j d_j,$$

其中 $x_1,x_2,\cdots,x_k$ 为(LP)的全部基本可行解,$d_1,d_2,\cdots,d_l$ 为 K 的全部极方向,且

$$\lambda_i \geqslant 0(i = 1,2,\cdots,k), \sum_{i=1}^{k} \lambda_i = 1, \mu_j \geqslant 0(j = 1,2,\cdots,l).$$

于是,(LP)等价于下面以 $\lambda_i \geqslant 0(i=1,2,\cdots,k)$ 和 $\mu_j \geqslant 0(j=1,2,\cdots,l)$ 为决策变量的线性规划问题

$$\begin{cases} \min & \sum_{i=1}^{k} (\boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}_i)\lambda_i + \sum_{j=1}^{l} (\boldsymbol{c}^{\mathrm{T}}\boldsymbol{d}_j)\mu_j; \\ \text{s.t.} & \sum_{i=1}^{k} \lambda_i = 1, \\ & \lambda_i \geqslant 0, i = 1,2,\cdots,k, \\ & \mu_j \geqslant 0, j = 1,2,\cdots,l. \end{cases} \tag{4.1.5}$$

由于 μ_j 可以任意大，因此若存在某个 $\boldsymbol{d}_j$，使 $\boldsymbol{c}^{\mathrm{T}}\boldsymbol{d}_j<0$，则问题(4.1.5)的目标函数无下界，即问题(4.1.5)不存在最优解，从而(LP)不存在最优解.

若对一切 $j=1,2,\cdots,l$，均有 $\boldsymbol{c}^{\mathrm{T}}\boldsymbol{d}_j \geqslant 0$，设

$$\boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}_s = \min_{1\leqslant i\leqslant k} \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}_i,$$

则

$$\boldsymbol{c}^{\mathrm{T}}\boldsymbol{x} = \sum_{i=1}^{k} (\boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}_i)\lambda_i + \sum_{j=1}^{l} (\boldsymbol{c}^{\mathrm{T}}\boldsymbol{d}_j)\mu_j \geqslant \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}_s, \forall \boldsymbol{x} \in K,$$

即知基本可行解 $\boldsymbol{x}_s$ 是(LP)的最优解. □

由定理 4.1.5 及其证明，可以得到

推论 4.1.6 若(LP)的可行域 K 无界，且(LP)存在最优解，则至少存在一个基本可行解为最优解. □

定理 4.1.3、定理 4.1.4 和推论 4.1.6 统称为线性规划的基本定理. 基本定理告诉我们，如果(LP)有最优解，则可限制在基本可行解(即可行域的极点)中去寻找最优解.

最后我们给出(LP)的最优解的表示定理和(LP)的 K－T 条件.

定理 4.1.7 设在(LP)的全部基本可行解 $\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_k$ 中，使目标函数值最小者为 $\boldsymbol{x}_{i_1},\boldsymbol{x}_{i_2},\cdots,\boldsymbol{x}_{i_s}$；在 K 的全部极方向 $\boldsymbol{d}_1,\boldsymbol{d}_2,\cdots,\boldsymbol{d}_l$ 中，满足 $\boldsymbol{c}^{\mathrm{T}}\boldsymbol{d}_j=0$ 者为 $\boldsymbol{d}_{j_1}$，$\boldsymbol{d}_{j_2},\cdots,\boldsymbol{d}_{j_t}$. 若(LP)存在最优解，则 $\boldsymbol{x}$ 为(LP)的最优解的充要条件是存在 $\lambda_{i_p} \geqslant 0$ $(p=1,2,\cdots,s)$，$\sum_{p=1}^{s} \lambda_{i_p}=1$，$\mu_{j_q} \geqslant 0 (q=1,2,\cdots,t)$，使

$$\boldsymbol{x} = \sum_{p=1}^{s} \lambda_{i_p}\boldsymbol{x}_{i_p} + \sum_{q=1}^{t} \mu_{j_q}\boldsymbol{d}_{j_q}. \tag{4.1.6}$$

证明 因为(LP)存在最优解，所以由定理 4.1.4 和推论 4.1.6 及其证明知，基本可行解 $\boldsymbol{x}_{i_p}(p=1,2,\cdots,s)$ 是(LP)的最优解.

设 $\boldsymbol{x}$ 具有(4.1.6)式的形式，则由推论 2.2.6 和定理 2.2.10 知，$\boldsymbol{x}$ 为 (LP) 的可行解，从而由(4.1.6)式知

$$\boldsymbol{c}^{\mathrm{T}}\boldsymbol{x} = \sum_{p=1}^{s} \lambda_{i_p}\boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}_{i_p} + \sum_{q=1}^{t} \mu_{j_p}\boldsymbol{c}^{\mathrm{T}}\boldsymbol{d}_{j_q} = \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}_{i_1},$$

因此,$\boldsymbol{x}$ 为(LP)的最优解.

反之,设 $\boldsymbol{x}$ 为(LP)的任一最优解,则 $\boldsymbol{x}$ 为可行解,于是由推论 2.2.6 和定理 2.2.10 知,存在 $\lambda_i \geqslant 0 (i=1,2,\cdots,k)$, $\sum_{i=1}^{k} \lambda_i = 1, \mu_j \geqslant 0 (j=1,2,\cdots,l)$,使

$$\boldsymbol{x} = \sum_{i=1}^{k} \lambda_i \boldsymbol{x}_i + \sum_{j=1}^{l} \mu_j \boldsymbol{d}_j. \tag{4.1.7}$$

根据定理 4.1.5,有

$$\boldsymbol{c}^{\mathrm{T}} \boldsymbol{d}_j \geqslant 0, j = 1,2,\cdots,l, \tag{4.1.8}$$

且由 $\boldsymbol{x}_{i_1}$ 为最优解知

$$\boldsymbol{c}^{\mathrm{T}} \boldsymbol{x}_i \geqslant \boldsymbol{c}^{\mathrm{T}} \boldsymbol{x}_{i_1}, i = 1,2,\cdots,k. \tag{4.1.9}$$

从而由(4.1.8)和(4.1.9)两式容易用反证法证明:若(4.1.7)式中某个 $\lambda_i > 0$,则 $\boldsymbol{x}_i$ 必为(LP)的最优解;若(4.1.7)式中某个 $\mu_j > 0$,则必有 $\boldsymbol{c}^{\mathrm{T}} \boldsymbol{d}_j = 0$. 由此知,最优解 $\boldsymbol{x}$ 必具有(4.1.6)式的形式. □

这个定理表明,若(LP)有两个最优解,则它必有无穷多个最优解. 因此(LP)的解有四种可能:

(1) (LP)有惟一最优解(此时,$\boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}$ 的最优值恰在 K 的一个极点上达到);

(2) (LP)有无穷多个最优解(此时,$\boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}$ 的最优值在 K 的一条棱上达到);

(3) (LP)有可行解,但无最优解(此时,K 无界且 $\boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}$ 在 K 上无下界);

(4) (LP)无可行解.

定理 4.1.8 设 $\bar{\boldsymbol{x}}$ 为(LP)的可行解,则 $\bar{\boldsymbol{x}}$ 为(LP)的最优解的充要条件是存在 $\boldsymbol{y} \in \mathbb{R}^m$,使得

$$\begin{cases} \boldsymbol{A}^{\mathrm{T}} \boldsymbol{y} \leqslant \boldsymbol{c}, \\ (\boldsymbol{c} - \boldsymbol{A}^{\mathrm{T}} \boldsymbol{y})^{\mathrm{T}} \bar{\boldsymbol{x}} = 0. \end{cases} \tag{4.1.10}$$

证明 充分性. 设 $\bar{\boldsymbol{x}}$ 为(LP)的可行解,且存在 $\boldsymbol{y} \in \mathbb{R}^m$,使得(4.1.10)式成立,则对于(LP)的任一可行解 $\boldsymbol{x}$,有

$$\boldsymbol{c}^{\mathrm{T}} \boldsymbol{x} \geqslant \boldsymbol{y}^{\mathrm{T}} \boldsymbol{A} \boldsymbol{x} = \boldsymbol{y}^{\mathrm{T}} \boldsymbol{b} = \boldsymbol{y}^{\mathrm{T}} \boldsymbol{A} \bar{\boldsymbol{x}} = \boldsymbol{c}^{\mathrm{T}} \bar{\boldsymbol{x}},$$

这说明 $\bar{\boldsymbol{x}}$ 为(LP)的最优解.

必要性. 设 $\bar{\boldsymbol{x}}$ 是(LP)的最优解,不妨设 $\bar{x}_j > 0 (j=1,2,\cdots,k)$, $\bar{x}_{k+1} = \bar{x}_{k+2} = \cdots = \bar{x}_n = 0$. 由此可知关系式组

$$\begin{cases} \boldsymbol{A}\boldsymbol{u} \leqslant \boldsymbol{0}, \\ -\boldsymbol{A}\boldsymbol{u} \leqslant \boldsymbol{0}, \\ u_j \geqslant 0 (j = k+1, k+2, \cdots, n), \\ \boldsymbol{c}^{\mathrm{T}} \boldsymbol{u} < 0 \end{cases} \tag{4.1.11}$$

无解,其中 $\boldsymbol{u} = (u_1, u_2, \cdots, u_n)^{\mathrm{T}}$. 这是因为,若不然,可取充分小的 $\lambda > 0$,使得 $\bar{\boldsymbol{x}} + \lambda \boldsymbol{u} \geqslant \boldsymbol{0}$,且

$$A(\bar{x}+\lambda u)=A\bar{x}+\lambda Au=b,$$

即 $\bar{x}+\lambda u$ 为(LP)的可行解,但是

$$c^{\mathrm{T}}(\bar{x}+\lambda u)=c^{\mathrm{T}}\bar{x}+\lambda c^{\mathrm{T}}u<c^{\mathrm{T}}\bar{x},$$

此与 $\bar{x}$ 是(LP)的最优解相矛盾.

若记

$$\widetilde{A}=(A^{\mathrm{T}},-A^{\mathrm{T}},-e_{k+1},-e_{k+2},\cdots,-e_n)^{\mathrm{T}},$$

这里 e_j 是第 j 个分量为1其余分量为0的 n 维向量,$j=k+1,k+2,\cdots,n$,则关系式组(4.1.11)无解等价于关系式组

$$\widetilde{A}u\leqslant 0,\ -c^{\mathrm{T}}u>0$$

无解,根据 Farkas 引理,这又等价于关系式组

$$\widetilde{A}^{\mathrm{T}}v=-c,v\geqslant 0$$

有解. 再记 $v=(v_1^{\mathrm{T}},v_2^{\mathrm{T}},v_{k+1},v_{k+2},\cdots,v_n)^{\mathrm{T}}$,其中 $v_1\in\mathbb{R}^m,v_2\in\mathbb{R}^m,v_j\in\mathbb{R}(j=k+1,k+2,\cdots,n)$,于是有 $v_1\geqslant 0,v_2\geqslant 0,v_j\geqslant 0(j=k+1,k+2,\cdots,n)$,且

$$A^{\mathrm{T}}(v_1-v_2)-(0,0,\cdots,0,v_{k+1},v_{k+2},\cdots,v_n)^{\mathrm{T}}=-c,$$

令 $y=v_2-v_1$,则

$$c-A^{\mathrm{T}}y=(0,0,\cdots,0,v_{k+1},v_{k+2},\cdots,v_n)^{\mathrm{T}}.$$

由于 $v_j\geqslant 0,\bar{x}_j=0,j=k+1,k+2,\cdots,n$,因此(4.1.10)式成立. □

与3.3节和3.4节一样,称(4.1.10)式为(LP)的K-T条件.

4.2 单纯形法

在4.1节我们已经指出:求解(LP),只需在基本可行解中搜索(如果(LP)存在最优解的话). 然而求出并比较所有的基本可行解的枚举方法通常是不实用的,因为尽管基本可行解的个数不会超过$\binom{n}{m}$,但当 n 较大时,$\binom{n}{m}$将是一个非常大的数!

Dantzig 在1947年提出的单纯形法(simplex method)是一种行之有效的方法. 它的基本思想是:从一个基本可行解出发,求出使目标函数值下降的另一个基本可行解;通过不断改进基本可行解,力图找出使目标函数值达到最小的基本可行解. 为实现这种基本思想,我们需要解决下面三个问题:

(1) 求(LP)的初始基本可行解的方法;

(2) 判别一个基本可行解是否为最优解的准则;

(3) 从一个基本可行解转换到使目标函数值下降的另一个基本可行解的方法.

下面我们分别来讨论这些问题.

4.2.1 最优性条件

设 $\bar{x}$ 是(LP)的一个基本可行解,为了叙述上的方便,先设 $\bar{x}$ 对应的基为

$$B = (p_1, p_2, \cdots, p_m),$$

从而 $x_1, x_2, \cdots, x_m$ 为基变量,$x_{m+1}, x_{m+2}, \cdots, x_n$ 为非基变量. 记

$$N = (p_{m+1}, p_{m+2}, \cdots, p_n),$$

$$x_B = (x_1, x_2, \cdots, x_m)^{\mathrm{T}},$$

$$x_N = (x_{m+1}, x_{m+2}, \cdots, x_n)^{\mathrm{T}}.$$

于是

$$A = (B, N), x = \begin{bmatrix} x_B \\ x_N \end{bmatrix}, c = \begin{bmatrix} c_B \\ c_N \end{bmatrix},$$

即知 $Ax = b$ 等价于

$$Bx_B + Nx_N = b.$$

由此解得

$$x_B = B^{-1}b - B^{-1}Nx_N, \tag{4.2.1}$$

这是用非基变量表示基变量的公式.

在(4.2.1)式中,令 $x_N = 0$,即知 $x_B = B^{-1}b$,从而得基本解

$$\bar{x} = \begin{bmatrix} B^{-1}b \\ 0 \end{bmatrix}.$$

将(4.2.1)式代入目标函数,有

$$\begin{aligned} c^{\mathrm{T}}x &= c_B^{\mathrm{T}}x_B + c_N^{\mathrm{T}}x_N \\ &= c_B^{\mathrm{T}}(B^{-1}b - B^{-1}Nx_N) + c_N^{\mathrm{T}}x_N \\ &= c_B^{\mathrm{T}}B^{-1}b - (c_B^{\mathrm{T}}B^{-1}N - c_N^{\mathrm{T}})x_N, \end{aligned}$$

即得用非基变量表示目标函数的公式:

$$f = c_B^{\mathrm{T}}B^{-1}b - (c_B^{\mathrm{T}}B^{-1}N - c_N^{\mathrm{T}})x_N.$$

以上推导表明,对于给定的一个基 B,(LP)可化为如下的等价形式:

$$\begin{cases} \min & f = c_B^{\mathrm{T}}B^{-1}b - (c_B^{\mathrm{T}}B^{-1}N - c_N^{\mathrm{T}})x_N; \\ \text{s.t.} & x_B + B^{-1}Nx_N = B^{-1}b, \\ & x \geqslant 0. \end{cases} \tag{4.2.2}$$

称(4.2.2)式为(LP)关于基 B(或基本可行解 $\bar{x}$)的典式(canonical form).

如果 $\bar{x}$ 对应的基 B 为一般形式,即

$$B = (p_{j_1}, p_{j_2}, \cdots, p_{j_m}),$$

则通过类似的推导,可得关于一般基 B 的典式仍具有(4.2.2)式的形式. 只是此时,$x_B = (x_{j_1}, x_{j_2}, \cdots, x_{j_m})^{\mathrm{T}}$,$x_N$ 为非基变量构成的 $n-m$ 维向量;N 是非基变量对

应的列向量构成的 $m\times(n-m)$ 矩阵;$\boldsymbol{c}_B=(c_{j_1},c_{j_2},\cdots,c_{j_m})^{\mathrm{T}}$,$\boldsymbol{c}_N$ 为目标函数中非基变量的系数构成的 $n-m$ 维向量.

下面把关于一般基 $\boldsymbol{B}$ 的典式(4.2.2)用代数式来表示.

记 $R=\{1,2,\cdots,n\}\setminus\{j_1,j_2,\cdots,j_m\}$,它表示非基变量的指标集,并令

$$\boldsymbol{B}^{-1}\boldsymbol{b}=\begin{bmatrix}b_{10}\\b_{20}\\\vdots\\b_{m0}\end{bmatrix},\boldsymbol{B}^{-1}\boldsymbol{A}=\begin{bmatrix}b_{11}&b_{12}&\cdots&b_{1n}\\b_{21}&b_{22}&\cdots&b_{2n}\\\vdots&\vdots&&\vdots\\b_{m1}&b_{m2}&\cdots&b_{mn}\end{bmatrix},$$

$$\pi_j=\boldsymbol{c}_B^{\mathrm{T}}\boldsymbol{B}^{-1}\boldsymbol{p}_j-c_j,\ \forall j\in R,$$

$$\bar{f}=\boldsymbol{c}_B^{\mathrm{T}}\boldsymbol{B}^{-1}\boldsymbol{b},$$

则(4.2.2)式等价于

$$\begin{cases}\min\quad f=\bar{f}-\sum\limits_{j\in R}\pi_jx_j;\\ \text{s.t.}\quad x_{j_i}+\sum\limits_{j\in R}b_{ij}x_j=b_{i0},i=1,2,\cdots,m,\\ \qquad\quad x_j\geqslant0,j=1,2,\cdots,n.\end{cases}\tag{4.2.3}$$

记

$$\boldsymbol{\pi}=(\boldsymbol{c}_B^{\mathrm{T}}\boldsymbol{B}^{-1}\boldsymbol{A}-\boldsymbol{c}^{\mathrm{T}})^{\mathrm{T}},$$

则基变量对应的部分

$$\boldsymbol{\pi}_B=(\boldsymbol{c}_B^{\mathrm{T}}\boldsymbol{B}^{-1}\boldsymbol{B}-\boldsymbol{c}_B^{\mathrm{T}})^{\mathrm{T}}=\boldsymbol{0};$$

而非基变量对应的部分

$$\boldsymbol{\pi}_N=(\boldsymbol{c}_B^{\mathrm{T}}\boldsymbol{B}^{-1}\boldsymbol{N}-\boldsymbol{c}_N^{\mathrm{T}})^{\mathrm{T}},$$

它是由前面所定义的 $\pi_j(j\in R)$ 构成的向量.

定理 4.2.1 设 $\bar{\boldsymbol{x}}$ 是(LP)的关于 $\boldsymbol{B}$ 的基本可行解,若 $\boldsymbol{\pi}_N\leqslant\boldsymbol{0}$,则 $\bar{\boldsymbol{x}}$ 是(LP)的最优解;若 $\bar{\boldsymbol{x}}$ 是(LP)的非退化的最优解,则 $\boldsymbol{\pi}_N\leqslant\boldsymbol{0}$.

证明 设 $\boldsymbol{\pi}_N\leqslant\boldsymbol{0}$,则由(4.2.2)式知,对一切 $\boldsymbol{x}\in K$,有

$$\begin{aligned}\boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}&=\boldsymbol{c}_B^{\mathrm{T}}\boldsymbol{B}^{-1}\boldsymbol{b}-(\boldsymbol{c}_B^{\mathrm{T}}\boldsymbol{B}^{-1}\boldsymbol{N}-\boldsymbol{c}_N^{\mathrm{T}})\boldsymbol{x}_N\\&\geqslant\boldsymbol{c}_B^{\mathrm{T}}\boldsymbol{B}^{-1}\boldsymbol{b}=\boldsymbol{c}^{\mathrm{T}}\bar{\boldsymbol{x}},\end{aligned}$$

所以 $\bar{\boldsymbol{x}}$ 是(LP)的最优解.

设 $\bar{\boldsymbol{x}}$ 是(LP)的非退化的最优解,则由定理 4.1.8 知,$\bar{\boldsymbol{x}}$ 满足 K-T 条件

$$\begin{cases}\boldsymbol{c}-\boldsymbol{A}^{\mathrm{T}}\boldsymbol{y}\geqslant\boldsymbol{0},\\(\boldsymbol{c}-\boldsymbol{A}^{\mathrm{T}}\boldsymbol{y})^{\mathrm{T}}\bar{\boldsymbol{x}}=0,\end{cases}$$

从而

$$\boldsymbol{c}_N-\boldsymbol{N}^{\mathrm{T}}\boldsymbol{y}\geqslant\boldsymbol{0},\tag{4.2.4}$$

$$(\boldsymbol{c}_B-\boldsymbol{B}^{\mathrm{T}}\boldsymbol{y})^{\mathrm{T}}\bar{\boldsymbol{x}}_B=0.\tag{4.2.5}$$

注意到 $\bar{x}_B = B^{-1}b > 0$,所以由(4.2.5)式有

$$c_B - B^{\mathrm{T}}y = 0.$$

即得 $y = (B^{-1})^{\mathrm{T}}c_B$,代入(4.2.4)式得

$$c_N - N^{\mathrm{T}}(B^{-1})^{\mathrm{T}}c_B \geq 0.$$

因此 $\pi_N = (c_B^{\mathrm{T}}B^{-1}N - c_N^{\mathrm{T}})^{\mathrm{T}} \leq 0.$ □

由此,我们把在最优性检验中起决定作用的系数 π_j 称为变量 x_j 的检验数(criterion).

定理 4.2.2 设 B 是(LP)的一个基,若关于 B 的典式(4.2.3)中存在 $r \in R$,使

$$b_{ir} \leq 0, i = 1,2,\cdots,m, \tag{4.2.6}$$

则存在可行域 K 的一个极方向 d,使 $c^{\mathrm{T}}d = -\pi_r$.

证明 令 $d = (d_1, d_2, \cdots, d_n)^{\mathrm{T}}$ 如下:

$$d_j = \begin{cases} -b_{ir}, j = j_i (i = 1,2,\cdots,m), \\ 1, \quad j = r, \\ 0, \quad j \in R\backslash\{r\}, \end{cases} \tag{4.2.7}$$

则 $d \geq 0, d \neq 0$,且由

$$p_r = B(B^{-1}p_r) = B\begin{bmatrix} b_{1r} \\ b_{2r} \\ \vdots \\ b_{mr} \end{bmatrix} = \sum_{i=1}^{m} b_{ir}p_{j_i} \tag{4.2.8}$$

有

$$Ad = \sum_{j=1}^{n} d_j p_j = p_r - \sum_{i=1}^{m} b_{ir}p_{j_i} = 0.$$

从而由定理 2.2.7 知,d 为 K 的一个方向.要证,d 为 K 的一个极方向,根据引理 2.2.8,只要证 $\bar{d} = \frac{1}{e^{\mathrm{T}}d} d$ 为

$$\bar{K}_0 = \{y \mid Ay = 0, e^{\mathrm{T}}y = 1, y \geq 0\}$$

的一个极点,容易知道,向量组

$$\bar{p}_{j_1} = \begin{bmatrix} p_{j_1} \\ 1 \end{bmatrix}, \cdots, \bar{p}_{j_m} = \begin{bmatrix} p_{j_m} \\ 1 \end{bmatrix}, \bar{p}_r = \begin{bmatrix} p_r \\ 1 \end{bmatrix}$$

线性无关,假若它们线性相关,则由 $p_{j_1}, p_{j_2}, \cdots, p_{j_m}$ 线性无关知 $\bar{p}_{j_1}, \bar{p}_{j_2}, \cdots, \bar{p}_{j_m}$ 线性无关,从而 $\bar{p}_r$ 可由这 m 个向量线性表出,即存在数 $\alpha_1, \alpha_2, \cdots, \alpha_m$,使 $\bar{p}_r = \sum_{i=1}^{m} \alpha_i \bar{p}_{j_i}$. 因此有

$$\boldsymbol{p}_r = \sum_{i=1}^{m} \alpha_i \boldsymbol{p}_{j_i}, \quad 1 = \sum_{i=1}^{m} \alpha_i. \tag{4.2.9}$$

另一方面，由于 $\boldsymbol{p}_{j_1}, \boldsymbol{p}_{j_2}, \cdots, \boldsymbol{p}_{j_m}$ 线性无关，因此由(4.2.8)和(4.2.9)两式知 $b_{ir} = \alpha_i (i = 1, 2, \cdots, m)$，从而 $\sum_{i=1}^{m} b_{ir} = 1$，此与 $b_{ir} \leqslant 0 (i = 1, 2, \cdots, m)$ 相矛盾. 根据定理 2.2.2，可知，$\bar{\boldsymbol{d}}$ 是 $\bar{K}_0$ 的极点. 显然还有

$$\boldsymbol{c}^{\mathrm{T}} \boldsymbol{d} = c_r - \sum_{i=1}^{m} c_{j_i} b_{ir} = c_r - \boldsymbol{c}_B^{\mathrm{T}} \boldsymbol{B}^{-1} \boldsymbol{p}_r = -\pi_r.$$

□

从定理 4.2.2 的证明可以看出，即使(4.2.6)式不成立，(4.2.7)式所定义的 $\boldsymbol{d}$ 仍满足 $\boldsymbol{A}\boldsymbol{d} = \boldsymbol{0}$，$\boldsymbol{c}^{\mathrm{T}}\boldsymbol{d} = -\pi_r$，但此时 $\boldsymbol{d}$ 不是可行域 K 的方向.

定理 4.2.3 设 $\bar{\boldsymbol{x}}$ 为(LP)的基本可行解，若关于 $\bar{\boldsymbol{x}}$ 的典式(4.2.3)中有某个检验数 $\pi_r > 0 (r \in R)$，且

$$b_{ir} \leqslant 0, i = 1, 2, \cdots, m,$$

则(LP)无最优解.

证明 由定理 4.2.2 知，存在可行域 K 的一个极方向 $\boldsymbol{d}$，使 $\boldsymbol{c}^{\mathrm{T}}\boldsymbol{d} = -\pi_r < 0$. 根据定理 4.1.5，(LP)无最优解. □

4.2.2 基本可行解改进

首先给出基的转换的一个结论.

定理 4.2.4 设 $\boldsymbol{B} = (\boldsymbol{p}_{j_1}, \boldsymbol{p}_{j_2}, \cdots, \boldsymbol{p}_{j_m})$ 是(LP)的一个基，且 $r \in R$，则

$$\widetilde{\boldsymbol{B}} = (\boldsymbol{p}_{j_1}, \cdots, \boldsymbol{p}_{j_{s-1}}, \boldsymbol{p}_r, \boldsymbol{p}_{j_{s+1}}, \cdots, \boldsymbol{p}_{j_m})$$

为(LP)的一个基的充要条件是在关于 $\boldsymbol{B}$ 的典式(4.2.3)中 $b_{sr} \neq 0$.

证明 若 $b_{sr} \neq 0$，则由(4.2.8)式知

$$\boldsymbol{p}_{j_s} = \frac{1}{b_{sr}} \boldsymbol{p}_r + \sum_{i \neq s} \left(-\frac{b_{ir}}{b_{sr}} \right) \boldsymbol{p}_{j_i},$$

这说明 $\boldsymbol{B}$ 的第 s 列 $\boldsymbol{p}_{j_s}$ 可由 $\widetilde{\boldsymbol{B}}$ 的各列线性表出，又显然 $\boldsymbol{B}$ 的其他列也都可由 $\widetilde{\boldsymbol{B}}$ 的各列线性表出，所以

$$m = \operatorname{rank} \boldsymbol{B} \leqslant \operatorname{rank} \widetilde{\boldsymbol{B}} \leqslant m,$$

因此 $\operatorname{rank} \widetilde{\boldsymbol{B}} = m$，即知 $\widetilde{\boldsymbol{B}}$ 为(LP)的一个基.

若 $b_{sr} = 0$，则由(4.2.8)式知

$$\boldsymbol{p}_r + \sum_{i \neq s} (-b_{ir}) \boldsymbol{p}_{j_i} = \boldsymbol{0},$$

即 $\widetilde{\boldsymbol{B}}$ 的列向量组是线性相关的，从而 $\widetilde{\boldsymbol{B}}$ 不是(LP)的基. □

现在考虑定理 4.2.1 和定理 4.2.3 均不满足的情形. 这时，我们有

定理 4.2.5 设 $\bar{\boldsymbol{x}}$ 为(LP)的非退化的基本可行解，若关于 $\bar{\boldsymbol{x}}$ 的典式(4.2.3)

中有 $\pi_r>0(r\in R)$，且至少有一个 $b_{ir}>0(1\leqslant i\leqslant m)$，则必存在另一个基本可行解 $\tilde{\boldsymbol{x}}$，使 $\boldsymbol{c}^{\mathrm{T}}\tilde{\boldsymbol{x}}<\boldsymbol{c}^{\mathrm{T}}\bar{\boldsymbol{x}}$.

证明 取

$$\tilde{\boldsymbol{x}} = \bar{\boldsymbol{x}} + \theta \boldsymbol{d}, \tag{4.2.10}$$

其中 $\boldsymbol{d}$ 由(4.2.7)式所定义，且

$$\theta = \min\left\{\frac{b_{i0}}{b_{ir}}\middle| b_{ir} > 0, i = 1,2,\cdots,m\right\} = \frac{b_{s0}}{b_{sr}}. \tag{4.2.11}$$

根据定理 4.2.2 后面的说明有 $\boldsymbol{A}\boldsymbol{d}=\boldsymbol{0}$，即知

$$\boldsymbol{A}\tilde{\boldsymbol{x}} = \boldsymbol{A}\bar{\boldsymbol{x}} + \theta\boldsymbol{A}\boldsymbol{d} = \boldsymbol{b}.$$

而 $\bar{\boldsymbol{x}}$ 是非退化的，即 $b_{i0}>0(i=1,2,\cdots,m)$，因此 $\theta>0$. 于是，当 $b_{ir}\leqslant 0$ 时，

$$b_{i0} - \theta b_{ir} \geqslant b_{i0} > 0;$$

当 $b_{ir}>0$ 时，

$$b_{i0} - \theta b_{ir} \geqslant b_{i0} - \frac{b_{i0}}{b_{ir}}b_{ir} = 0.$$

从而由 $\tilde{\boldsymbol{x}}$ 的定义知 $\tilde{x}_j\geqslant 0(j=1,2,\cdots,n)$，故 $\tilde{\boldsymbol{x}}$ 为(LP)的可行解.

又由(4.2.11)式知 $\tilde{x}_{j_s}=0$，即知 $\tilde{x}$ 的非零分量至多为

$$\tilde{x}_{j_1},\cdots,\tilde{x}_{j_{s-1}},\tilde{x}_r,\tilde{x}_{j_{s+1}},\cdots,\tilde{x}_{j_m},$$

且由 $b_{sr}>0$ 及定理 4.2.4 知

$$\widetilde{\boldsymbol{B}} = (\boldsymbol{p}_{j_1},\cdots,\boldsymbol{p}_{j_{s-1}},\boldsymbol{p}_r,\boldsymbol{p}_{j_{s+1}},\cdots,\boldsymbol{p}_{j_m})$$

为(LP)的一个基，所以根据定理 4.1.1，$\tilde{\boldsymbol{x}}$ 为(LP)的基本可行解. 并且由 $\theta>0$，$\pi_r>0$ 及定理 4.2.2 后面的说明可知

$$\boldsymbol{c}^{\mathrm{T}}\tilde{\boldsymbol{x}} = \boldsymbol{c}^{\mathrm{T}}\bar{\boldsymbol{x}} + \theta\boldsymbol{c}^{\mathrm{T}}\boldsymbol{d} = \boldsymbol{c}^{\mathrm{T}}\bar{\boldsymbol{x}} - \theta\pi_r < \boldsymbol{c}^{\mathrm{T}}\bar{\boldsymbol{x}}.$$ □

4.2.3 单纯形法的算法步骤

在定理 4.2.5 的证明中，实际上给出了改进基本可行解的一个方法：把对应于正检验数 π_r 的非基变量 x_r 变成基变量，称它为进基变量(incoming variable)，而从原基变量中按(4.2.11)式确定 x_{j_s} 变为非基变量，称它为离基变量(outgoing variable).

现在来讨论如何从关于基 $\boldsymbol{B}=(\boldsymbol{p}_{j_1},\boldsymbol{p}_{j_2},\cdots,\boldsymbol{p}_{j_m})$ 的典式(4.2.3)式导出关于新基 $\widetilde{\boldsymbol{B}}=(\boldsymbol{p}_{j_1},\cdots,\boldsymbol{p}_{j_{s-1}},\boldsymbol{p}_r,\boldsymbol{p}_{j_{s+1}},\cdots,\boldsymbol{p}_{j_m})$ 的典式. 为此将典式(4.2.3)中的系数写成表 4.2.1 的表格形式.

表 4.2.1　单纯形表 $T(\boldsymbol{B})$

	x_1	$\cdots$	x_r	$\cdots$	x_n	
x_{j_1}	b_{11}	$\cdots$	b_{1r}	$\cdots$	b_{1n}	b_{10}
$\vdots$	$\vdots$		$\vdots$		$\vdots$	$\vdots$
x_{j_s}	b_{s1}	$\cdots$	b_{sr}	$\cdots$	b_{sn}	b_{s0}
$\vdots$	$\vdots$		$\vdots$		$\vdots$	$\vdots$
x_{j_m}	b_{m1}	$\cdots$	b_{mr}	$\cdots$	b_{mn}	b_{m0}
f	π_1	$\cdots$	π_r	$\cdots$	π_n	$\bar{f}$

这个表称为(LP)关于基 $\boldsymbol{B}$ 的单纯形表(simplex tabuleau),记为 $T(\boldsymbol{B})$. 于是只要说明怎样从原单纯形表 $T(\boldsymbol{B})$ 导出新的单纯形表 $T(\tilde{\boldsymbol{B}})$ 即可. 按照解线性方程组的 Gauss - Jordan 消去法思想,要使 x_r 变成基变量,必须把 $T(\boldsymbol{B})$ 中 x_r 所在的列变成单位向量. 因此可得单纯形表的变换规则如下:

(1) 把 $T(\boldsymbol{B})$ 中第 s 行同除以 b_{sr} 作为新的第 s 行(这样把 x_r 所在的列中第 s 个元素变成 1),即

$$(s\text{行}) := \frac{1}{b_{sr}}(s\text{行});$$

(2) 把表中新的第 s 行乘以 $(-b_{ir})$ 加到第 i 行 $(i \neq s)$,得到新的第 i 行(把 x_r 所在的列中第 i 个元素变成 0),即

$$(i\text{行}) := (i\text{行}) - b_{ir}(s\text{行}),\ \forall i \in \{1,2,\cdots,m\} \setminus \{s\};$$

(3) 把表中新的第 s 行乘以 $(-\pi_r)$ 加到第 $m+1$ 行得到新的第 $m+1$ 行(把 x_r 的检验数变成 0),即

$$(m+1\text{行}) := (m+1\text{行}) - \pi_r(s\text{行}).$$

上述变换称为 $\{s,r\}$ 旋转变换(pivot transformation),元素 b_{sr} 称为主元(pivot element),主元所在的行和列分别称为主元行(pivot row)和主元列(pivot column).

算法 4-1(单纯形法)

Step 1　对于一个已知的可行基 $\boldsymbol{B} = (\boldsymbol{p}_{j_1}, \boldsymbol{p}_{j_2}, \cdots, \boldsymbol{p}_{j_m})$,求出关于 $\boldsymbol{B}$ 的单纯形表 $T(\boldsymbol{B})$.

Step 2　如果所有 $\pi_j \leqslant 0\ (j = 1,2,\cdots,n)$,则关于 $\boldsymbol{B}$ 的基本可行解 $\bar{\boldsymbol{x}}$ 便是(LP)的最优解,$\bar{f}$ 是最优值,此时的 $T(\boldsymbol{B})$ 称为最优单纯形表,算法结束;否则,转 Step 3.

Step 3　如果有 $\pi_r > 0$,使 $T(\boldsymbol{B})$ 中 x_r 所在的列

$$(b_{1r}, b_{2r}, \cdots, b_{mr})^{\mathrm{T}} \leqslant \boldsymbol{0},$$

则(LP)无最优解,算法终止;否则,转 Step 4.

Step 4　令 r 为最大正检验数中指标最小者,即

$$r = \min\{l \mid \pi_l = \max_{\pi_j > 0} \pi_j\}, \tag{4.2.12}$$

取 x_r 为进基变量;令 j_s 为比值最小的行中指标最小者,即

$$j_s = \min\left\{j_k \left| \frac{b_{k0}}{b_{kr}} = \min_{b_{ir} > 0} \frac{b_{i0}}{b_{ir}} \right.\right\}, \tag{4.2.13}$$

取 x_{j_s} 为离基变量.

Step 5　以 b_{sr} 为主元进行 $\{s,r\}$ 旋转变换,得到新的单纯形表 $T(\widetilde{\boldsymbol{B}})$,以 $\widetilde{\boldsymbol{B}}$ 取代 $\boldsymbol{B}$,返回 Step 2.

从 Step 2 到 Step 5 的每一次循环称为一次单纯形迭代(simplex iteration).

式(4.2.12)和(4.2.13)分别称为 Dantzig 进基规则(incoming rule)和离基规则(outgoing rule),统称为 Dantzig 规则.

例 4.2.1　求解线性规则问题

$$\begin{cases} \min \quad f = -2x_1 - x_2; \\ \text{s.t.} \quad x_1 + x_2 + x_3 = 5, \\ \qquad -x_1 + x_2 + x_4 = 0, \\ \qquad 6x_1 + 2x_2 + x_5 = 21, \\ \qquad x_j \geqslant 0, j = 1,2,3,4,5. \end{cases}$$

解　问题有一个明显的可行基 $\boldsymbol{B} = (\boldsymbol{p}_3, \boldsymbol{p}_4, \boldsymbol{p}_5)$,对应的初始基本可行解 $\bar{\boldsymbol{x}} = (0,0,5,0,21)^{\mathrm{T}}$,相应的单纯形表为表 4.2.2.

表 4.2.2　初始单纯形表

	x_1	x_2	x_3	x_4	x_5	
x_3	1	1	1	0	0	5
x_4	-1	1	0	1	0	0
x_5	6^*	2	0	0	1	21
f	2	1	0	0	0	0

因为 $\pi_1 = 2 > 0, \pi_2 = 1 > 0$,所以取 x_1 为进基变量. 而

$$\theta = \min\left\{\frac{5}{1}, \frac{21}{6}\right\} = \frac{21}{6} = \frac{b_{30}}{b_{31}},$$

故取 x_5 为离基变量,$b_{31} = 6$ 为主元(在主元右上角标以“*”号),进行 $\{3,1\}$ 旋转变换,得单纯形表 4.2.3.

表 4.2.3　第一轮迭代后的单纯形表

	x_1	x_2	x_3	x_4	x_5	
x_3	0	$2/3^*$	1	0	$-1/6$	3/2
x_4	0	4/3	0	1	1/6	7/2
x_1	1	1/3	0	0	1/6	7/2
f	0	1/3	0	0	$-1/3$	-7

同理,再进行{1,2}旋转变换,得单纯形表 4.2.4.

表 4.2.4　第二轮迭代后的最优单纯形表

	x_1	x_2	x_3	x_4	x_5	
x_2	0	1	3/2	0	$-1/4$	9/4
x_4	0	0	-2	1	1/2	1/2
x_1	1	0	$-1/2$	0	1/4	11/4
f	0	0	$-1/2$	0	$-1/4$	$-31/4$

此时所有检验数非正,得最优解

$$\bar{x} = \left(\frac{11}{4},\frac{9}{4},0,\frac{1}{2},0\right)^{\mathrm{T}},$$

最优解值为 $-31/4$. □

4.2.4　退化情形的处理

对于非退化的(LP),使用单纯形法经有限次迭代,必能求得(LP)的最优解或判定(LP)无最优解,这是因为由定理 4.2.5 知,每次迭代都使目标函数值严格单调减少,即知出现过的基本可行解不会再次出现,而基本可行解的个数有限. 对于退化的(LP),可以照常使用单纯形法,只要在迭代过程中基不重复出现(基本可行解可能重复出现),由于基的个数有限,因此经有限次迭代也能得出最优解或判定(LP)无最优解. 但如果基在迭代过程中重复出现,则会使以后的迭代成为前面迭代的重复,从而迭代会无限循环下去,算法无法结束. 这种现象称为基的循环(basic cycling). 1951 年,Hoffman 首先构造出一个出现基循环的例子. 1955 年,Beale 构造了一个更简单的例子如下:

$$
\begin{cases}
\min \quad f = -\dfrac{3}{4}x_4 + 20x_5 - \dfrac{1}{2}x_6 + 6x_7; \\
\text{s.t.} \quad x_1 + \dfrac{1}{4}x_4 - 8x_5 - x_6 + 9x_7 = 0, \\
\qquad x_2 + \dfrac{1}{2}x_4 - 12x_5 - \dfrac{1}{2}x_6 + 3x_7 = 0, \\
\qquad x_3 + x_6 = 1, \\
\qquad x_j \geqslant 0, j = 1,2,\cdots,7.
\end{cases}
$$

对 Beale 例子用单纯形法迭代 6 次，基发生循环（请读者自己验证）：

$$
(\boldsymbol{p}_1,\boldsymbol{p}_2,\boldsymbol{p}_3) \to (\boldsymbol{p}_4,\boldsymbol{p}_2,\boldsymbol{p}_3) \to (\boldsymbol{p}_4,\boldsymbol{p}_5,\boldsymbol{p}_3) \to (\boldsymbol{p}_6,\boldsymbol{p}_5,\boldsymbol{p}_3) \to
$$
$$
(\boldsymbol{p}_6,\boldsymbol{p}_7,\boldsymbol{p}_3) \to (\boldsymbol{p}_1,\boldsymbol{p}_7,\boldsymbol{p}_3) \to (\boldsymbol{p}_1,\boldsymbol{p}_2,\boldsymbol{p}_3).
$$

为了避免基的循环，1952 年，Charnes 提出了摄动法（perturbation method）；1954 年，Dantzig，Orden 和 Wolfe 提出了字典序规则（lexicographic rule）；但是最简便的还是 Bland 在 1977 年提出的如下的一个规则，称之为 Bland 规则.

（1）进基规则：由

$$
r = \min\{j \mid \pi_j > 0\}
$$

确定 x_r 为进基变量；

（2）离基规则（同 Dantzig 离基规则）：由

$$
j_s = \min\left\{ j_k \,\middle|\, \frac{b_{k0}}{b_{kr}} = \min_{b_{ir}>0} \frac{b_{i0}}{b_{ir}} \right\}
$$

确定 x_{j_s} 为离基变量.

利用 Bland 规则求解 Beale 例子，不会发生基的循环. 读者可以验证，经过 3 次单纯形迭代，得到最优解和最优解值如下：

$$
\bar{\boldsymbol{x}} = \left(\frac{3}{4},0,0,1,0,1,0\right)^{\mathrm{T}}, \min f = -\frac{5}{4}.
$$

可以证明：任何（LP）按 Bland 规则选取进基变量和离基变量，不会出现基的循环（参见文献[51]）.

从理论上看，在单纯形算法步骤中采用 Bland 规则，可以使算法在有限步结束. 但具体计算表明，在各种选主元规则中，Dantzig 规则的效果最好，即在求解同一线性规划问题时，迭代次数最少. 在实际计算中，退化是常有的，但发生基的循环却是罕见的（Hoffman 和 Beale 的例子都是人为的）. 因此，对于线性规划问题仍可按 Dantzig 规则进行单纯形迭代，万一遇到基的循环，再改用 Bland 规则.

4.2.5 初始基本可行解的求法

用单纯形法解（LP）时，需要先有一个初始基本可行解，如果像例 4.2.1 那

样，在约束矩阵中含有一个单位矩阵 $\boldsymbol{I}$，且 $\boldsymbol{b}\geqslant\boldsymbol{0}$，则可取初始可行基为 $\boldsymbol{B}=\boldsymbol{I}$，即可得到一个基本可行解. 但是在一般情况下，难以凭观察得出初始可行基，甚至有无可行基都难以判定. 因此有必要给出寻找初始基本可行解的一般方法.

仍考虑标准形式的线性规划问题(LP)

$$\begin{cases}\min & \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x};\\ \text{s.t.} & \boldsymbol{A}\boldsymbol{x}=\boldsymbol{b},\\ & \boldsymbol{x}\geqslant\boldsymbol{0},\end{cases}$$

且不妨设 $\boldsymbol{b}\geqslant\boldsymbol{0}$，但并不要求 $\boldsymbol{A}$ 为行满秩矩阵.

寻找初始基本可行解的方法主要有两阶段法(two phase method)和大 M 法(big M method). 下面我们只介绍两阶段法. 关于大 M 法可参见文献[52].

在第一阶段先求解如下的线性规划问题

$$\begin{cases}\min & g=\sum_{i=1}^{m}x_{n+i};\\ \text{s.t.} & \sum_{j=1}^{n}a_{ij}x_j+x_{n+i}=b_i,\ i=1,2,\cdots,m,\\ & x_j\geqslant 0,\ j=1,2,\cdots,n+m,\end{cases}\tag{LP1}$$

其中 $x_{n+i}(i=1,2,\cdots,m)$ 称为人工变量(artificial variables). 因 $\boldsymbol{b}\geqslant\boldsymbol{0}$，故(LP1)有一个现成的基本可行解：

$$x_j=0,\ j=1,2,\cdots,n;$$

$$x_{n+i}=b_i,\ i=1,2,\cdots,m,$$

与之对应的基为单位矩阵，从而目标函数 g 可改写为

$$g=\sum_{i=1}^{m}x_{n+i}=\sum_{i=1}^{m}\Big(b_i-\sum_{j=1}^{n}a_{ij}x_j\Big)=\sum_{i=1}^{m}b_i-\sum_{j=1}^{n}\sum_{i=1}^{m}a_{ij}x_j,$$

于是得到(LP1)的一个单纯形表如表 4.2.5.

表 4.2.5　两阶段法初始单纯形表

	x_1	$\cdots$	x_n	x_{n+1}	$\cdots$	x_{n+m}	
x_{n+1}	a_{11}	$\cdots$	a_{1n}	1	$\cdots$	0	b_1
$\vdots$	$\vdots$		$\vdots$	$\vdots$		$\vdots$	$\vdots$
x_{n+m}	a_{m1}	$\cdots$	a_{mn}	0	$\cdots$	1	b_m
g	$\sum_{i=1}^{m}a_{i1}$	$\cdots$	$\sum_{i=1}^{m}a_{in}$	0	$\cdots$	0	$\sum_{i=1}^{m}b_i$

由于目标函数 $g\geqslant 0$,即它在(LP1)的可行域上有下界,因此(LP1)必有最优解. 于是从单纯形表 4.2.5 出发,通过单纯形迭代必可求得(LP1)的最优解. 设最优解为$\begin{bmatrix}\hat{x}\\ \hat{y}\end{bmatrix}$,其中 $\hat{x}=(\hat{x}_1,\hat{x}_2,\cdots,\hat{x}_n)^{\mathrm{T}}$,$\hat{y}=(\hat{x}_{n+1},\hat{x}_{n+2},\cdots,\hat{x}_{n+m})^{\mathrm{T}}$,对应的基为 $\hat{B}$,分三种情况讨论.

(1) $\hat{y}\neq 0$. 此时(LP)无可行解. 因为假若(LP)存在一个可行解 x,则$\begin{bmatrix}x\\ 0\end{bmatrix}$为(LP1)的可行解,且对应的目标函数 g 的值为 0,这与 $\min g=\sum\limits_{i=1}^{m}\hat{x}_{n+i}>0$ 相矛盾.

(2) $\hat{y}=0$ 且人工变量 $x_{n+i}(i=1,2,\cdots,m)$都是非基变量. 这时 $\hat{x}$ 是(LP)的可行解. 又因基变量全在 $x_1,x_2,\cdots,x_n$ 之中,故对应的基 $\hat{B}$ 必为 A 的子方阵,所以 $\hat{x}$ 为(LP)的基本可行解.

(3) $\hat{y}=0$ 且基变量中含有人工变量,设 x_{n+s}为基变量,则(LP1)关于 $\hat{B}$ 的单纯形表 $T(\hat{B})$ 中 x_{n+s}所在的第 s 行对应的方程为

$$x_{n+s}+\sum_{j\in J}b_{sj}x_j+\sum_{i\in I}b_{s,n+i}x_{n+i}=0, \tag{4.2.14}$$

这里 J 为 $x_1,x_2,\cdots,x_n$ 中非基变量的指标集,I 为人工变量中非基变量的指标集.

如果(4.2.14)式中所有 $b_{sj}=0(j\in J)$,则有

$$x_{n+s}+\sum_{i\in I}b_{s,n+i}x_{n+i}=0.$$

这说明人工变量 x_{n+s}可由诸人工变量 $x_{n+i}(i\in I)$线性表出,从而可知原约束方程组 $Ax=b$ 中第 s 个方程可由另外一些方程(即人工变量 $x_{n+i}(i\in I)$对应的那些约束方程)的适当线性组合来表示,因此,第 s 个约束方程是多余的,应当删去,这相当于从 $T(\hat{B})$ 中删去第 s 行.

如果(4.2.14)中存在 $r\in J$,使 $b_{sr}\neq 0$,则由定理 4.2.4 知,以 b_{sr}为主元进行 $\{s,r\}$旋转变换,得到(LP1)的新的单纯形表,它对应的基本可行解仍为(LP1)的最优解,但新的基变量中减少了一个人工变量 x_{n+s}. 若新的基变量中还有人工变量,再重复此法,经过有限次,必能化为(2)的情形.

综上所述,对于不具有明显可行基的(LP),可先用单纯形法解(LP1),解的结果或者说明(LP)无可行解,或者找到(LP)的一个基本可行解. 然后再从这个基本可行解开始应用单纯形法求解(LP),这是两阶段法的第二阶段. 由此不难写出两阶段法的算法步骤,这留给读者去完成.

值得注意的是,在第一阶段迭代过程中,人工变量一旦离开基,随之也就失去了作用,故可从单纯形表中删去人工变量所在的列.

例 4.2.2 求解线性规划问题

$$
\begin{cases}
\min & f = -x_1 - 3x_2 + x_3; \\
\text{s. t.} & x_2 + 2x_3 + x_4 = 4, \\
& -x_1 + 2x_2 + x_3 + x_4 + x_5 = 4, \\
& 3x_1 + 3x_3 + x_4 = 4, \\
& x_j \geqslant 0, j = 1,2,\cdots,5.
\end{cases}
\tag{4.2.15}
$$

解 只需引进两个人工变量 x_6 和 x_7,相应的(LP1)如下:

$$
\begin{cases}
\min & g = x_6 + x_7; \\
\text{s. t.} & x_2 + 2x_3 + x_4 + x_6 = 4, \\
& -x_1 + 2x_2 + x_3 + x_4 + x_5 = 4, \\
& 3x_1 + 3x_3 + x_4 + x_7 = 4, \\
& x_j \geqslant 0, j = 1,2,\cdots,7.
\end{cases}
$$

取初始可行基 $\boldsymbol{B} = (\boldsymbol{p}_5, \boldsymbol{p}_6, \boldsymbol{p}_7)$,相应的单纯形表如表 4.2.6.

表 4.2.6 例 4.2.2 第一阶段初始单纯形表

	x_1	x_2	x_3	x_4	x_5	x_6	x_7	
x_6	0	1	2	1	0	1	0	4
x_5	-1	2	1	1	1	0	0	4
x_7	3	0	3^*	1	0	0	1	4
g	3	1	5	2	0	0	0	8
f	1	3	-1	0	0	0	0	0

在表 4.2.6 中添加原问题(4.2.15)的目标函数 f 对应的行是为过渡到第二阶段作准备. 用单纯形法进行迭代的过程见表 4.2.7 和表 4.2.8(因人工变量 x_7 离基,故在表中删去 x_7 所在的列).

表 4.2.7 例 4.2.2 第一阶段迭代一次后的单纯形表

	x_1	x_2	x_3	x_4	x_5	x_6	
x_6	-2	1	0	1/3	0	1	4/3
x_5	-2	2^*	0	2/3	1	0	8/3
x_3	1	0	1	1/3	0	0	4/3
g	-2	1	0	1/3	0	0	4/3
f	2	3	0	1/3	0	0	4/3

表 4.2.8　例 4.2.2 第一阶段迭代两次后的最优单纯形表

	x_1	x_2	x_3	x_4	x_5	x_6	
x_6	-1^*	0	0	0	$-1/2$	1	0
x_2	-1	1	0	1/3	1/2	0	4/3
x_3	1	0	1	1/3	0	0	4/3
g	-1	0	0	0	$-1/2$	0	0
f	5	0	0	$-2/3$	$-3/2$	0	$-8/3$

至此得到了(LP1)的最优解,但是 x_6 为基变量且 $b_{11}\neq 0$,于是在表 4.2.8 中以 b_{11} 为主元作旋转变换,得表 4.2.9(同理,在表中删去已经离基的人工变量 x_6 所在的列).

表 4.2.9　例 4.2.2 第一阶段最优单纯形表删去人工变量后的单纯形表

	x_1	x_2	x_3	x_4	x_5	
x_1	1	0	0	0	1/2	0
x_2	0	1	0	1/3	1	4/3
x_3	0	0	1	1/3	$-1/2$	4/3
g	0	0	0	0	0	0
f	0	0	0	$-2/3$	-4	$-8/3$

因此第一阶段结束,把表 4.2.9 中 g 行删除,便得问题(4.2.15)的初始单纯形表:

表 4.2.10　例 4.2.2 第二阶段初始单纯形表即最优单纯形表

	x_1	x_2	x_3	x_4	x_5	
x_1	1	0	0	0	1/2	0
x_2	0	1	0	1/3	1	4/3
x_3	0	0	1	1/3	$-1/2$	4/3
f	0	0	0	$-2/3$	-4	$-8/3$

因为表 4.2.10 中的检验数全部非正,所以得到问题(4.2.15)的最优解和最优值分别为

$$\tilde{\boldsymbol{x}} = \left(0,\frac{4}{3},\frac{4}{3},0,0\right)^{\mathrm{T}},\tilde{f} = -\frac{8}{3}.$$ □

4.2.6 单纯形法的几何解释

我们考虑在一次单纯形迭代中新旧基本可行解之间的几何特性.

定理 4.2.6 设 $\bar{\boldsymbol{x}}$ 是(LP)关于基 $\boldsymbol{B}$ 的基本可行解,对 $T(\boldsymbol{B})$ 进行一次单纯形迭代得到新的基本可行解 $\tilde{\boldsymbol{x}}$,若 $\bar{\boldsymbol{x}}$ 是非退化的,则 $\bar{\boldsymbol{x}}$ 与 $\tilde{\boldsymbol{x}}$ 是(LP)的可行域 K 的相邻极点.

证明 设关于 $\boldsymbol{B}$ 的典式中的约束方程组为

$$x_{j_i} = \bar{x}_{j_i} - \sum_{j \in R} b_{ij}x_j, i = 1,2,\cdots,m, \tag{4.2.16}$$

并设迭代时,进基变量为 x_r,离基变量为 x_{j_s},则由单纯形迭代规则可知

$$\tilde{\boldsymbol{x}} = \bar{\boldsymbol{x}} + \theta\boldsymbol{d} = \bar{\boldsymbol{x}} + \tilde{x}_r\boldsymbol{d}, \tag{4.2.17}$$

其中 $\boldsymbol{d}$ 按(4.2.7)式定义,θ 按(4.2.11)式定义.

因 $\bar{\boldsymbol{x}}$ 是非退化的,故 $\theta>0$,从而 $\tilde{\boldsymbol{x}}\neq\bar{\boldsymbol{x}}$,且 $\boldsymbol{d}$ 与 $\tilde{\boldsymbol{x}}-\bar{\boldsymbol{x}}$ 的指向相同. 所以对任意 $\boldsymbol{x}_0 \in \overline{\bar{\boldsymbol{x}}\tilde{\boldsymbol{x}}}$,有

$$\boldsymbol{x}_0 = \bar{\boldsymbol{x}} + \beta\boldsymbol{d}, 0 \leqslant \beta \leqslant \theta. \tag{4.2.18}$$

若存在 $\boldsymbol{x}_1,\boldsymbol{x}_2 \in K$,使

$$\boldsymbol{x}_0 = \alpha\boldsymbol{x}_1 + (1-\alpha)\boldsymbol{x}_2, 0 < \alpha < 1, \tag{4.2.19}$$

往证 $\boldsymbol{x}_1,\boldsymbol{x}_2 \in \overline{\bar{\boldsymbol{x}}\tilde{\boldsymbol{x}}}$.

由 $\bar{x}_j=\tilde{x}_j=0(j\in R\setminus\{r\})$ 和(4.2.18)式知

$$x_j^{(0)} = 0, \forall j \in R\setminus\{r\},$$

从而由(4.2.19)式有

$$x_j^{(1)} = x_j^{(2)} = 0, \forall j \in R\setminus\{r\}.$$

又因为 $\boldsymbol{A}\boldsymbol{x}_1=\boldsymbol{b},\boldsymbol{A}\boldsymbol{x}_2=\boldsymbol{b}$,所以由(4.2.16)式和 $\boldsymbol{d}$ 的定义便知

$$\boldsymbol{x}_1 = \bar{\boldsymbol{x}} + x_r^{(1)}\boldsymbol{d}, \boldsymbol{x}_2 = \bar{\boldsymbol{x}} + x_r^{(2)}\boldsymbol{d}.$$

记 $\gamma_k=x_r^{(k)}/\theta(k=1,2)$,则由(4.2.17)式有

$$\boldsymbol{x}_k = \bar{\boldsymbol{x}} + \gamma_k(\tilde{\boldsymbol{x}}-\bar{\boldsymbol{x}}) = (1-\gamma_k)\bar{\boldsymbol{x}} + \gamma_k\tilde{\boldsymbol{x}}, k = 1,2. \tag{4.2.20}$$

若 $\gamma_k>1$,则

$$\tilde{\boldsymbol{x}} = \frac{1}{\gamma_k}\boldsymbol{x}_k + \left(1-\frac{1}{\gamma_k}\right)\bar{\boldsymbol{x}},$$

由 $\tilde{\boldsymbol{x}}\neq\bar{\boldsymbol{x}}$ 知此与 $\tilde{\boldsymbol{x}}$ 为极点相矛盾. 因此 $0\leqslant\gamma_k\leqslant1(k=1,2)$,根据(4.2.20)式,$\boldsymbol{x}_1$,$\boldsymbol{x}_2 \in \overline{\bar{\boldsymbol{x}}\tilde{\boldsymbol{x}}}$. 这就证明了 $\bar{\boldsymbol{x}}$ 与 $\tilde{\boldsymbol{x}}$ 是(LP)的可行域 K 的相邻极点. □

根据定理 4.2.6,单纯形法的几何意义就是从(LP)的可行域 K 的一个极点 $\bar{\boldsymbol{x}}$ 出发,沿多胞形 K 的棱求出与 $\bar{\boldsymbol{x}}$ 相邻并使目标函数值下降的另一个极点,如此

继续下去，直至得出最优解或判定问题无最优解.

在非退化的假设下，因为 $n-m$ 个非基变量的每一个都能去替代给定的基本可行解中的一个基变量，所以可行域 K 中任何一个极点都有 $n-m$ 个极点与它相邻，而这 $n-m+1$ 个极点正好构成(LP)的约束方程组解空间$\mathbb{R}^{n-m}$中的一个单纯形，这就是把算法 4－1 称为单纯形法的原因.

4.3 对偶理论

随着线性规划问题及其解法的提出，人们发现，对于任何一个线性规划问题，都伴随着另一个线性规划问题，二者之间存在密切的关系，因而常常把这两个问题放在一起研究. 这就形成了线性规划中重要而有趣的对偶理论.

定义 4.3.1 设有线性规划问题

$$\begin{cases}\min & \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x};\\ \text{s.t.} & \boldsymbol{A}\boldsymbol{x}\geqslant\boldsymbol{b},\\ & \boldsymbol{x}\geqslant\boldsymbol{0},\end{cases}\tag{4.3.1}$$

其中

$$\boldsymbol{x}=\begin{bmatrix}x_1\\x_2\\\vdots\\x_n\end{bmatrix},\boldsymbol{c}=\begin{bmatrix}c_1\\c_2\\\vdots\\c_n\end{bmatrix},\boldsymbol{b}=\begin{bmatrix}b_1\\b_2\\\vdots\\b_m\end{bmatrix},\boldsymbol{A}=\begin{bmatrix}a_{11}&a_{12}&\cdots&a_{1n}\\a_{21}&a_{22}&\cdots&a_{2n}\\\vdots&\vdots&&\vdots\\a_{m1}&a_{m2}&\cdots&a_{mn}\end{bmatrix}.$$

考虑另一个线性规划问题

$$\begin{cases}\max & \boldsymbol{b}^{\mathrm{T}}\boldsymbol{y};\\ \text{s.t.} & \boldsymbol{A}^{\mathrm{T}}\boldsymbol{y}\leqslant\boldsymbol{c},\\ & \boldsymbol{y}\geqslant\boldsymbol{0},\end{cases}\tag{4.3.2}$$

这里 $\boldsymbol{y}=(y_1,y_2,\cdots,y_m)^{\mathrm{T}}$，我们称问题(4.3.2)为问题(4.3.1)的对偶问题(dual problem)，并称问题(4.3.1)为原问题(prime problem).

下面用一个实例来说明对偶线性规划问题的意义.

例 4.3.1(营养问题) 某饲养场所用的饲料由 n 种配料混合而成. 要求这种饲料必须含有 m 种营养成分，而且每单位饲料中第 i 种营养成分的含量不能低于 b_i. 已知第 i 种营养成分在每单位第 j 种配料中的含量为 a_{ij}，第 j 种配料的单位价格为 c_j. 问在保证营养要求的条件下，应采用何种配方才能使饲料的费用最小?

解 设 x_j 为每单位饲料中第 j 种配料的含量($j=1,2,\cdots,n$)，则营养问题的数学模型为

$$
\begin{cases}
\min & \sum_{j=1}^{n} c_j x_j; \\
\text{s.t.} & \sum_{j=1}^{n} a_{ij} x_j \geqslant b_i, i = 1,2,\cdots,m, \\
& x_j \geqslant 0, j = 1,2,\cdots,n.
\end{cases}
\tag{4.3.3}
$$

现在从另一个角度提出如下问题：某饲料公司欲把这 m 种营养成分分别制成 m 种营养丸出售. 为了使饲养场能采用公司生产的营养丸替代天然配料，就必须做到营养丸的价格不超过与之相当的天然配料的价格. 公司面临的问题是，在上述条件的限制下，如何确定各种营养丸的单位价格，才能使公司获利最大？

设第 i 种营养丸的单价为 $y_i(i=1,2,\cdots,m)$，则 $a_{ij}y_i$ 表示把单位第 j 种配料中第 i 种营养成分折合成营养丸的代价，于是这个问题的数学模型为

$$
\begin{cases}
\max & \sum_{i=1}^{m} b_i y_i; \\
\text{s.t.} & \sum_{i=1}^{m} a_{ij} y_i \leqslant c_j, j = 1,2,\cdots,n, \\
& y_i \geqslant 0, i = 1,2,\cdots,m,
\end{cases}
\tag{4.3.4}
$$

这里问题(4.3.4)就是问题(4.3.3)的对偶问题. 这说明，原问题和对偶问题是同一件事情的两个不同的侧面，它们之间有着密切的依存关系. □

定理 4.3.1 对偶问题的对偶是原问题.

证明 将对偶问题(4.3.2)化成与原问题(4.3.1)相同的形式：

$$
\begin{cases}
\min & (-\boldsymbol{b})^{\mathrm{T}}\boldsymbol{y}; \\
\text{s.t.} & (-\boldsymbol{A})^{\mathrm{T}}\boldsymbol{y} \geqslant -\boldsymbol{c}, \\
& \boldsymbol{y} \geqslant \boldsymbol{0},
\end{cases}
\tag{4.3.5}
$$

由定义 4.3.1 知，问题(4.3.5)的对偶问题为

$$
\begin{cases}
\max & (-\boldsymbol{c})^{\mathrm{T}}\boldsymbol{x}; \\
\text{s.t.} & -\boldsymbol{A}\boldsymbol{x} \leqslant -\boldsymbol{b}, \\
& \boldsymbol{x} \geqslant \boldsymbol{0},
\end{cases}
\tag{4.3.6}
$$

而问题(4.3.6)等价于问题(4.3.1). 这表明，问题(4.3.2)的对偶问题为问题(4.3.1). □

因此，定义 4.3.1 中所确定的对偶关系是互相的，即原问题与对偶问题的地位是可以相互转换的.

标准形式的线性规划问题

$$\begin{cases}\min & \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x};\\ \text{s.t.} & \boldsymbol{A}\boldsymbol{x}=\boldsymbol{b},\\ & \boldsymbol{x}\geqslant \boldsymbol{0},\end{cases}\tag{LP}$$

可以改写为问题(4.3.1)的形式

$$\begin{cases}\min & \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x};\\ \text{s.t.} & \begin{bmatrix}\boldsymbol{A}\\ -\boldsymbol{A}\end{bmatrix}\boldsymbol{x}\geqslant\begin{bmatrix}\boldsymbol{b}\\ -\boldsymbol{b}\end{bmatrix},\\ & \boldsymbol{x}\geqslant \boldsymbol{0}.\end{cases}$$

从而它的对偶问题为

$$\begin{cases}\max & (\boldsymbol{b}^{\mathrm{T}},-\boldsymbol{b}^{\mathrm{T}})\boldsymbol{w};\\ \text{s.t.} & (\boldsymbol{A}^{\mathrm{T}},-\boldsymbol{A}^{\mathrm{T}})\boldsymbol{w}\leqslant \boldsymbol{c},\\ & \boldsymbol{w}\geqslant \boldsymbol{0}.\end{cases}\tag{4.3.7}$$

再令 $\boldsymbol{w}=\begin{bmatrix}\boldsymbol{w}_1\\ \boldsymbol{w}_2\end{bmatrix}$,则问题(4.3.7)等价于

$$\begin{cases}\max & \boldsymbol{b}^{\mathrm{T}}(\boldsymbol{w}_1-\boldsymbol{w}_2);\\ \text{s.t.} & \boldsymbol{A}^{\mathrm{T}}(\boldsymbol{w}_1-\boldsymbol{w}_2)\leqslant \boldsymbol{c},\\ & \boldsymbol{w}_1\geqslant \boldsymbol{0},\boldsymbol{w}_2\geqslant \boldsymbol{0}.\end{cases}\tag{4.3.8}$$

记 $\boldsymbol{y}=\boldsymbol{w}_1-\boldsymbol{w}_2$,由问题(4.3.8),得到(LP)的对偶问题如下

$$\begin{cases}\max & \boldsymbol{b}^{\mathrm{T}}\boldsymbol{y};\\ \text{s.t.} & \boldsymbol{A}^{\mathrm{T}}\boldsymbol{y}\leqslant \boldsymbol{c}.\end{cases}\tag{DP}$$

因为任何一个线性规划问题都可化成(LP)的形式,所以任一线性规划问题都有对偶问题.

对于一般的线性规划问题

$$\begin{cases}\min & \boldsymbol{c}_1^{\mathrm{T}}\boldsymbol{x}_1+\boldsymbol{c}_2^{\mathrm{T}}\boldsymbol{x}_2;\\ \text{s.t.} & \boldsymbol{A}_{11}\boldsymbol{x}_1+\boldsymbol{A}_{12}\boldsymbol{x}_2\geqslant \boldsymbol{b}_1,\\ & \boldsymbol{A}_{21}\boldsymbol{x}_1+\boldsymbol{A}_{22}\boldsymbol{x}_2=\boldsymbol{b}_2,\\ & \boldsymbol{x}_1\geqslant \boldsymbol{0},\end{cases}\tag{P}$$

其中 $\boldsymbol{A}_{ij}$ 是 $m_i\times n_j$ 矩阵,$\boldsymbol{b}_i$ 为 m_i 维向量,$\boldsymbol{c}_j$ 为 n_j 维向量,$\boldsymbol{x}_j$ 为 n_j 维向量,$i=1,2$,$j=1,2$.

引入 n_1 维向量 $\boldsymbol{w}\geqslant\boldsymbol{0}$,且令 $\boldsymbol{x}_2=\boldsymbol{x}_{21}-\boldsymbol{x}_{22}$,$\boldsymbol{x}_{21}\geqslant\boldsymbol{0}$,$\boldsymbol{x}_{22}\geqslant\boldsymbol{0}$,则(P)可以化为标准形式

$$\begin{cases}\min & \boldsymbol{c}_1^{\mathrm{T}}\boldsymbol{x}_1+\boldsymbol{c}_2^{\mathrm{T}}\boldsymbol{x}_{21}-\boldsymbol{c}_2^{\mathrm{T}}\boldsymbol{x}_{22};\\ \text{s.t.} & \boldsymbol{A}_{11}\boldsymbol{x}_1+\boldsymbol{A}_{12}\boldsymbol{x}_{21}-\boldsymbol{A}_{12}\boldsymbol{x}_{22}-\boldsymbol{w}=\boldsymbol{b}_1,\\ & \boldsymbol{A}_{21}\boldsymbol{x}_1+\boldsymbol{A}_{22}\boldsymbol{x}_{21}-\boldsymbol{A}_{22}\boldsymbol{x}_{22}=\boldsymbol{b}_2,\\ & \boldsymbol{x}_1\geqslant\boldsymbol{0},\boldsymbol{x}_{21}\geqslant\boldsymbol{0},\boldsymbol{x}_{22}\geqslant\boldsymbol{0},\boldsymbol{w}\geqslant\boldsymbol{0}.\end{cases}$$

由前知,它的对偶问题为

$$\begin{cases}\max & \boldsymbol{b}_1^{\mathrm{T}}\boldsymbol{y}_1+\boldsymbol{b}_2^{\mathrm{T}}\boldsymbol{y}_2;\\ \text{s.t.} & \boldsymbol{A}_{11}^{\mathrm{T}}\boldsymbol{y}_1+\boldsymbol{A}_{21}^{\mathrm{T}}\boldsymbol{y}_2\leqslant\boldsymbol{c}_1,\\ & \boldsymbol{A}_{12}^{\mathrm{T}}\boldsymbol{y}_1+\boldsymbol{A}_{22}^{\mathrm{T}}\boldsymbol{y}_2\leqslant\boldsymbol{c}_2,\\ & -\boldsymbol{A}_{12}^{\mathrm{T}}\boldsymbol{y}_1-\boldsymbol{A}_{22}^{\mathrm{T}}\boldsymbol{y}_2\leqslant-\boldsymbol{c}_2,\\ & -\boldsymbol{y}_1\leqslant\boldsymbol{0}.\end{cases}$$

这等价于

$$\begin{cases}\max & \boldsymbol{b}_1^{\mathrm{T}}\boldsymbol{y}_1+\boldsymbol{b}_2^{\mathrm{T}}\boldsymbol{y}_2;\\ \text{s.t.} & \boldsymbol{A}_{11}^{\mathrm{T}}\boldsymbol{y}_1+\boldsymbol{A}_{21}^{\mathrm{T}}\boldsymbol{y}_2\leqslant\boldsymbol{c}_1,\\ & \boldsymbol{A}_{12}^{\mathrm{T}}\boldsymbol{y}_1+\boldsymbol{A}_{22}^{\mathrm{T}}\boldsymbol{y}_2=\boldsymbol{c}_2,\\ & \boldsymbol{y}_1\geqslant\boldsymbol{0}.\end{cases}\tag{D}$$

这表明,(P)的对偶问题为(D).

下面用表格形式给出由原问题构造对偶问题的一般方法.

表 4.3.1 原问题与对偶问题的对应关系

原问题	min	对偶问题	max
变	≥0	行	≤
		约	
量	无限制	束	=
行	≥	变	≥0
约			
束	=	量	无限制

这里,我们把约束条件分为行约束(变量的线性组合的等式或不等式约束)和变量的符号约束两部分,而以原问题的行约束与对偶问题的变量一一对应,原问题的变量与对偶问题的行约束一一对应.

例 4.3.2 求如下线性规划问题的对偶问题

$$\begin{cases}\min & 3x_1+2x_2-3x_3+4x_4;\\ \text{s.t.} & x_1-2x_2+3x_3+4x_4\leqslant 3,\\ & x_2+3x_3+4x_4\geqslant -5,\\ & 2x_1-3x_2-7x_3-4x_4=2,\\ & x_1\geqslant 0,x_4\leqslant 0.\end{cases}$$

解 先把上述线性规划问题写成如下形式

$$\begin{cases}\min & 3x_1+2x_2-3x_3-4x_4';\\ \text{s.t.} & -x_1+2x_2-3x_3+4x_4'\geqslant -3,\\ & x_2+3x_3-4x_4'\geqslant -5,\\ & 2x_1-3x_2-7x_3+4x_4'=2,\\ & x_1\geqslant 0,x_4'\geqslant 0.\end{cases}$$

它的对偶问题为

$$\begin{cases}\max & -3y_1-5y_2+2y_3;\\ \text{s.t.} & -y_1+2y_3\leqslant 3,\\ & 2y_1+y_2-3y_3=2,\\ & -3y_1+3y_2-7y_3=-3,\\ & y_1-y_2+y_3\leqslant -1,\\ & y_1\geqslant 0,y_2\geqslant 0.\end{cases}$$

□

下面我们来考察互为对偶的线性规划问题之间的内在联系.

定理 4.3.2 设$\begin{bmatrix}\boldsymbol{x}_1\\ \boldsymbol{x}_2\end{bmatrix}$和$\begin{bmatrix}\boldsymbol{y}_1\\ \boldsymbol{y}_2\end{bmatrix}$分别为(P)和(D)的可行解,则

$$\boldsymbol{c}_1^{\mathrm{T}}\boldsymbol{x}_1+\boldsymbol{c}_2^{\mathrm{T}}\boldsymbol{x}_2\geqslant \boldsymbol{b}_1^{\mathrm{T}}\boldsymbol{y}_1+\boldsymbol{b}_2^{\mathrm{T}}\boldsymbol{y}_2. \tag{4.3.9}$$

证明 因为 $\boldsymbol{x}_1\geqslant\boldsymbol{0},\boldsymbol{y}_1\geqslant\boldsymbol{0}$,并且

$$\boldsymbol{A}_{11}\boldsymbol{x}_1+\boldsymbol{A}_{12}\boldsymbol{x}_2\geqslant \boldsymbol{b}_1,\boldsymbol{A}_{21}\boldsymbol{x}_1+\boldsymbol{A}_{22}\boldsymbol{x}_2=\boldsymbol{b}_2,$$
$$\boldsymbol{A}_{11}^{\mathrm{T}}\boldsymbol{y}_1+\boldsymbol{A}_{21}^{\mathrm{T}}\boldsymbol{y}_2\leqslant \boldsymbol{c}_1,\boldsymbol{A}_{12}^{\mathrm{T}}\boldsymbol{y}_1+\boldsymbol{A}_{22}^{\mathrm{T}}\boldsymbol{y}_2=\boldsymbol{c}_2,$$

所以

$$\begin{aligned}\boldsymbol{c}_1^{\mathrm{T}}\boldsymbol{x}_1+\boldsymbol{c}_2^{\mathrm{T}}\boldsymbol{x}_2&\geqslant(\boldsymbol{A}_{11}^{\mathrm{T}}\boldsymbol{y}_1+\boldsymbol{A}_{21}^{\mathrm{T}}\boldsymbol{y}_2)^{\mathrm{T}}\boldsymbol{x}_1+(\boldsymbol{A}_{12}^{\mathrm{T}}\boldsymbol{y}_1+\boldsymbol{A}_{22}^{\mathrm{T}}\boldsymbol{y}_2)^{\mathrm{T}}\boldsymbol{x}_2\\ &=\boldsymbol{y}_1^{\mathrm{T}}(\boldsymbol{A}_{11}\boldsymbol{x}_1+\boldsymbol{A}_{12}\boldsymbol{x}_2)+\boldsymbol{y}_2^{\mathrm{T}}(\boldsymbol{A}_{21}\boldsymbol{x}_1+\boldsymbol{A}_{22}\boldsymbol{x}_2)\\ &\geqslant \boldsymbol{y}_1^{\mathrm{T}}\boldsymbol{b}_1+\boldsymbol{y}_2^{\mathrm{T}}\boldsymbol{b}_2.\end{aligned}$$

因此,(4.3.9)式成立. □

定理 4.3.3 (P)和(D)都有最优解的充要条件是它们都有可行解.

证明 必要性是显然的,只需证明充分性.

设(P)和(D)分别有可行解$\begin{bmatrix}\bar{\boldsymbol{x}}_1\\ \bar{\boldsymbol{x}}_2\end{bmatrix}$和$\begin{bmatrix}\bar{\boldsymbol{y}}_1\\ \bar{\boldsymbol{y}}_2\end{bmatrix}$,则由定理 4.3.2 知,对于(P)的任一可行解$\begin{bmatrix}\boldsymbol{x}_1\\ \boldsymbol{x}_2\end{bmatrix}$,都有

$$\boldsymbol{c}_1^{\mathrm{T}}\boldsymbol{x}_1+\boldsymbol{c}_2^{\mathrm{T}}\boldsymbol{x}_2\geqslant \boldsymbol{b}_1^{\mathrm{T}}\bar{\boldsymbol{y}}_1+\boldsymbol{b}_2^{\mathrm{T}}\bar{\boldsymbol{y}}_2,$$

即目标函数 $\boldsymbol{c}_1^{\mathrm{T}}\boldsymbol{x}_1+\boldsymbol{c}_2^{\mathrm{T}}\boldsymbol{x}_2$ 在(P)的可行域上有下界,所以(P)有最优解.

同样,对于(D)的任一可行解$\begin{bmatrix}\boldsymbol{y}_1\\ \boldsymbol{y}_2\end{bmatrix}$,都有

$$\boldsymbol{b}_1^{\mathrm{T}}\boldsymbol{y}_1+\boldsymbol{b}_2^{\mathrm{T}}\boldsymbol{y}_2\leqslant \boldsymbol{c}_1^{\mathrm{T}}\bar{\boldsymbol{x}}_1+\boldsymbol{c}_2^{\mathrm{T}}\bar{\boldsymbol{x}}_2,$$

即目标函数 $\boldsymbol{b}_1^{\mathrm{T}}\boldsymbol{y}_1+\boldsymbol{b}_2^{\mathrm{T}}\boldsymbol{y}_2$ 在(D)的可行域上有上界,所以(D)有最优解. □

定理 4.3.4 设$\begin{bmatrix}\bar{\boldsymbol{x}}_1\\ \bar{\boldsymbol{x}}_2\end{bmatrix}$为(P)的可行解,则它为(P)的最优解的充要条件是存在 $\boldsymbol{y}_1\in\mathbb{R}^{m_1}$,$\boldsymbol{y}_2\in\mathbb{R}^{m_2}$,使得

$$\begin{cases}\boldsymbol{A}_{11}^{\mathrm{T}}\boldsymbol{y}_1+\boldsymbol{A}_{21}^{\mathrm{T}}\boldsymbol{y}_2\leqslant \boldsymbol{c}_1,\\ \boldsymbol{A}_{12}^{\mathrm{T}}\boldsymbol{y}_1+\boldsymbol{A}_{22}^{\mathrm{T}}\boldsymbol{y}_2=\boldsymbol{c}_2,\\ \boldsymbol{y}_1\geqslant\boldsymbol{0},\\ \boldsymbol{y}_1^{\mathrm{T}}(\boldsymbol{A}_{11}\bar{\boldsymbol{x}}_1+\boldsymbol{A}_{12}\bar{\boldsymbol{x}}_2-\boldsymbol{b}_1)=0,\\ (\boldsymbol{c}_1-\boldsymbol{A}_{11}^{\mathrm{T}}\boldsymbol{y}_1+\boldsymbol{A}_{21}^{\mathrm{T}}\boldsymbol{y}_2)^{\mathrm{T}}\bar{\boldsymbol{x}}_1=0.\end{cases}$$

上式称为(P)的 K-T 条件.

证明 先把(P)化为标准形(这与求(P)的对偶问题(D)的过程相同),再应用定理 4.1.8 即得(P)的 K-T 条件. □

定理 4.3.5 设$\begin{bmatrix}\bar{\boldsymbol{x}}_1\\ \bar{\boldsymbol{x}}_2\end{bmatrix}$和$\begin{bmatrix}\bar{\boldsymbol{y}}_1\\ \bar{\boldsymbol{y}}_2\end{bmatrix}$分别为(P)和(D)的可行解,则它们分别为(P)和(D)的最优解的充要条件是

$$\boldsymbol{c}_1^{\mathrm{T}}\bar{\boldsymbol{x}}_1+\boldsymbol{c}_2^{\mathrm{T}}\bar{\boldsymbol{x}}_2=\boldsymbol{b}_1^{\mathrm{T}}\bar{\boldsymbol{y}}_1+\boldsymbol{b}_2^{\mathrm{T}}\bar{\boldsymbol{y}}_2. \tag{4.3.10}$$

证明 充分性. 设$\begin{bmatrix}\boldsymbol{x}_1\\ \boldsymbol{x}_2\end{bmatrix}$和$\begin{bmatrix}\boldsymbol{y}_1\\ \boldsymbol{y}_2\end{bmatrix}$分别为(P)和(D)的任一可行解,则由定理 4.3.2,有

$$\boldsymbol{c}_1^{\mathrm{T}}\boldsymbol{x}_1+\boldsymbol{c}_2^{\mathrm{T}}\boldsymbol{x}_2\geqslant \boldsymbol{b}_1^{\mathrm{T}}\bar{\boldsymbol{y}}_1+\boldsymbol{b}_2^{\mathrm{T}}\bar{\boldsymbol{y}}_2=\boldsymbol{c}_1^{\mathrm{T}}\bar{\boldsymbol{x}}_1+\boldsymbol{c}_2^{\mathrm{T}}\bar{\boldsymbol{x}}_2,$$
$$\boldsymbol{b}_1^{\mathrm{T}}\boldsymbol{y}_1+\boldsymbol{b}_2^{\mathrm{T}}\boldsymbol{y}_2\leqslant \boldsymbol{c}_1^{\mathrm{T}}\bar{\boldsymbol{x}}_1+\boldsymbol{c}_2^{\mathrm{T}}\bar{\boldsymbol{x}}_2=\boldsymbol{b}_1^{\mathrm{T}}\bar{\boldsymbol{y}}_1+\boldsymbol{b}_2^{\mathrm{T}}\bar{\boldsymbol{y}}_2.$$

因此,$\begin{bmatrix}\bar{\boldsymbol{x}}_1\\ \bar{\boldsymbol{x}}_2\end{bmatrix}$和$\begin{bmatrix}\bar{\boldsymbol{y}}_1\\ \bar{\boldsymbol{y}}_2\end{bmatrix}$分别是(P)和(D)的最优解.

必要性. 因$\begin{bmatrix}\bar{\boldsymbol{x}}_1\\ \bar{\boldsymbol{x}}_2\end{bmatrix}$是(P)的最优解,故由定理 4.3.4 知,下面的 K-T 条件

成立：

$$\begin{cases}\boldsymbol{A}_{11}^{\mathrm{T}}\boldsymbol{y}_1+\boldsymbol{A}_{21}^{\mathrm{T}}\boldsymbol{y}_2\leqslant \boldsymbol{c}_1,\\ \boldsymbol{A}_{12}^{\mathrm{T}}\boldsymbol{y}_1+\boldsymbol{A}_{22}^{\mathrm{T}}\boldsymbol{y}_2=\boldsymbol{c}_2,\\ \boldsymbol{y}_1^{\mathrm{T}}(\boldsymbol{A}_{11}\bar{\boldsymbol{x}}_1+\boldsymbol{A}_{12}\bar{\boldsymbol{x}}_2-\boldsymbol{b}_1)=0,\\ (\boldsymbol{c}_1-\boldsymbol{A}_{11}^{\mathrm{T}}\boldsymbol{y}_1-\boldsymbol{A}_{21}^{\mathrm{T}}\boldsymbol{y}_2)^{\mathrm{T}}\bar{\boldsymbol{x}}_1=0,\\ \boldsymbol{y}_1\geqslant \boldsymbol{0}.\end{cases}$$

从而$\begin{bmatrix}\boldsymbol{y}_1\\ \boldsymbol{y}_2\end{bmatrix}$是(D)的可行解，并且

$$\begin{aligned}\boldsymbol{c}_1^{\mathrm{T}}\bar{\boldsymbol{x}}_1+\boldsymbol{c}_2^{\mathrm{T}}\bar{\boldsymbol{x}}_2&=(\boldsymbol{A}_{11}^{\mathrm{T}}\boldsymbol{y}_1+\boldsymbol{A}_{21}^{\mathrm{T}}\boldsymbol{y}_2)^{\mathrm{T}}\bar{\boldsymbol{x}}_1+(\boldsymbol{A}_{12}^{\mathrm{T}}\boldsymbol{y}_1+\boldsymbol{A}_{22}^{\mathrm{T}}\boldsymbol{y}_2)^{\mathrm{T}}\bar{\boldsymbol{x}}_2\\ &=\boldsymbol{y}_1^{\mathrm{T}}(\boldsymbol{A}_{11}\bar{\boldsymbol{x}}_1+\boldsymbol{A}_{12}\bar{\boldsymbol{x}}_2)+\boldsymbol{y}_2^{\mathrm{T}}(\boldsymbol{A}_{21}\bar{\boldsymbol{x}}_1+\boldsymbol{A}_{22}\bar{\boldsymbol{x}}_2)\\ &=\boldsymbol{y}_1^{\mathrm{T}}\boldsymbol{b}_1+\boldsymbol{y}_2^{\mathrm{T}}\boldsymbol{b}_2.\end{aligned}$$

因此，由充分性知，$\begin{bmatrix}\boldsymbol{y}_1\\ \boldsymbol{y}_2\end{bmatrix}$是(D)的最优解. 再注意到$\begin{bmatrix}\bar{\boldsymbol{y}}_1\\ \bar{\boldsymbol{y}}_2\end{bmatrix}$也是(D)的最优解，所以

$$\boldsymbol{b}_1^{\mathrm{T}}\bar{\boldsymbol{y}}_1+\boldsymbol{b}_2^{\mathrm{T}}\bar{\boldsymbol{y}}_2=\boldsymbol{b}_1^{\mathrm{T}}\boldsymbol{y}_1+\boldsymbol{b}_2^{\mathrm{T}}\boldsymbol{y}_2=\boldsymbol{c}_1^{\mathrm{T}}\bar{\boldsymbol{x}}_1+\boldsymbol{c}_2^{\mathrm{T}}\bar{\boldsymbol{x}}_2.$$ □

定理 4.3.6 在(P)和(D)中，若一个有最优解，则另一个也有最优解，且(P)和(D)的最优值相等. 若其中一个有可行解但无最优解，则另一个必无可行解.

证明 由定理 4.3.5 的必要性证明可知，若(P)存在最优解，则(D)也有最优解，且(P)和(D)有相同的最优值. 又因为(P)和(D)是互为对偶的，所以若(D)有最优解，则(P)也有最优解.

根据定理 4.3.3，在(P)和(D)中，若其中一个有可行解但无最优解，则另一个必无可行解. □

定理 4.3.7 设$\begin{bmatrix}\bar{\boldsymbol{x}}_1\\ \bar{\boldsymbol{x}}_2\end{bmatrix}$和$\begin{bmatrix}\bar{\boldsymbol{y}}_1\\ \bar{\boldsymbol{y}}_2\end{bmatrix}$分别是(P)和(D)的可行解，则它们分别为(P)和(D)的最优解的充要条件是

$$\begin{cases}\bar{\boldsymbol{y}}_1^{\mathrm{T}}(\boldsymbol{A}_{11}\bar{\boldsymbol{x}}_1+\boldsymbol{A}_{12}\bar{\boldsymbol{x}}_2-\boldsymbol{b}_1)=0,\\ (\boldsymbol{c}_1-\boldsymbol{A}_{11}^{\mathrm{T}}\bar{\boldsymbol{y}}_1-\boldsymbol{A}_{21}^{\mathrm{T}}\bar{\boldsymbol{y}}_2)^{\mathrm{T}}\bar{\boldsymbol{x}}_1=0.\end{cases}\tag{4.3.11}$$

证明 因为$\begin{bmatrix}\bar{\boldsymbol{x}}_1\\ \bar{\boldsymbol{x}}_2\end{bmatrix}$和$\begin{bmatrix}\bar{\boldsymbol{y}}_1\\ \bar{\boldsymbol{y}}_2\end{bmatrix}$分别为(P)和(D)的可行解，所以由定理 4.3.5 知，它们分别为(P)和(D)的最优解的充要条件是(4.3.10)式成立.

由于

$$\boldsymbol{A}_{11}\bar{\boldsymbol{x}}_1+\boldsymbol{A}_{12}\bar{\boldsymbol{x}}_2\geqslant \boldsymbol{b}_1,\boldsymbol{A}_{21}\bar{\boldsymbol{x}}_1+\boldsymbol{A}_{22}\bar{\boldsymbol{x}}_2=\boldsymbol{b}_2,\bar{\boldsymbol{x}}_1\geqslant \boldsymbol{0},$$

$$\boldsymbol{A}_{11}^{\mathrm{T}}\bar{\boldsymbol{y}}_1+\boldsymbol{A}_{21}^{\mathrm{T}}\bar{\boldsymbol{y}}_2\leqslant \boldsymbol{c}_1,\boldsymbol{A}_{12}^{\mathrm{T}}\bar{\boldsymbol{y}}_1+\boldsymbol{A}_{22}^{\mathrm{T}}\bar{\boldsymbol{y}}_2=\boldsymbol{c}_2,\bar{\boldsymbol{y}}_1\geqslant \boldsymbol{0},$$

因此有

$$
\begin{aligned}
\boldsymbol{c}_1^{\mathrm{T}}\bar{\boldsymbol{x}}_1 + \boldsymbol{c}_2^{\mathrm{T}}\bar{\boldsymbol{x}}_2 &\geqslant (\boldsymbol{A}_{11}^{\mathrm{T}}\bar{\boldsymbol{y}}_1 + \boldsymbol{A}_{21}^{\mathrm{T}}\bar{\boldsymbol{y}}_2)^{\mathrm{T}}\bar{\boldsymbol{x}}_1 + (\boldsymbol{A}_{12}^{\mathrm{T}}\bar{\boldsymbol{y}}_1 + \boldsymbol{A}_{22}^{\mathrm{T}}\bar{\boldsymbol{y}}_2)^{\mathrm{T}}\bar{\boldsymbol{x}}_2 \\
&= \bar{\boldsymbol{y}}_1^{\mathrm{T}}(\boldsymbol{A}_{11}\bar{\boldsymbol{x}}_1 + \boldsymbol{A}_{12}\bar{\boldsymbol{x}}_2) + \bar{\boldsymbol{y}}_2^{\mathrm{T}}(\boldsymbol{A}_{21}\bar{\boldsymbol{x}}_1 + \boldsymbol{A}_{22}\bar{\boldsymbol{x}}_2) \\
&\geqslant \bar{\boldsymbol{y}}_1^{\mathrm{T}}\boldsymbol{b}_1 + \bar{\boldsymbol{y}}_2^{\mathrm{T}}\boldsymbol{b}_2.
\end{aligned}
$$

由此可知,(4.3.10)式成立的充要条件是上式中不等号全为等号,这又等价于

$$
\begin{cases}
\boldsymbol{c}_1^{\mathrm{T}}\bar{\boldsymbol{x}}_1 = (\boldsymbol{A}_{11}^{\mathrm{T}}\bar{\boldsymbol{y}}_1 + \boldsymbol{A}_{21}^{\mathrm{T}}\bar{\boldsymbol{y}}_2)^{\mathrm{T}}\bar{\boldsymbol{x}}_1, \\
\bar{\boldsymbol{y}}_1^{\mathrm{T}}(\boldsymbol{A}_{11}\bar{\boldsymbol{x}}_1 + \boldsymbol{A}_{12}\bar{\boldsymbol{x}}_2) = \bar{\boldsymbol{y}}_1^{\mathrm{T}}\boldsymbol{b}_1.
\end{cases} \tag{4.3.12}
$$

式(4.3.12)成立的充要条件是(4.3.11)式成立. □

式(4.3.11)就是(P)的最优解必须满足的 K－T 条件中的互补松弛条件. 它表明,对于(P)和(D)的最优解,若(P)的第 i 个不等式约束为松约束,则(D)的第 i 个非负约束为紧约束;若(D)的第 i 个非负约束为松约束,则(P)的第 i 个不等式约束为紧约束;若(P)的第 j 个非负约束为松约束,则(D)的第 j 个不等式约束为紧约束;若(D)的第 j 个不等式约束为松约束,则(P)的第 j 个非负约束为紧约束.

作为定理 4.3.7 的特殊情形,立即得到下面的两个常用的互补松弛条件.

推论 4.3.8 设 $\bar{\boldsymbol{x}}$ 和 $\bar{\boldsymbol{y}}$ 分别为原问题(4.3.1)和对偶问题(4.3.2)的可行解,则它们分别为问题(4.3.1)和(4.3.2)的最优解的充要条件是

$$
\begin{cases}
\bar{\boldsymbol{y}}^{\mathrm{T}}(\boldsymbol{A}\bar{\boldsymbol{x}} - \boldsymbol{b}) = 0, \\
(\boldsymbol{c} - \boldsymbol{A}^{\mathrm{T}}\bar{\boldsymbol{y}})^{\mathrm{T}}\bar{\boldsymbol{x}} = 0.
\end{cases}
$$

□

推论 4.3.9 设 $\bar{\boldsymbol{x}}$ 和 $\bar{\boldsymbol{y}}$ 分别为(LP)和(DP)的可行解,则它们分别为(LP)和(DP)的最优解的充要条件是

$$
(\boldsymbol{c} - \boldsymbol{A}^{\mathrm{T}}\bar{\boldsymbol{y}})^{\mathrm{T}}\bar{\boldsymbol{x}} = 0.
$$

□

例 4.3.3 利用互补松弛条件解线性规划问题

$$
\begin{cases}
\max & 4x_1 + 3x_2 + 6x_3; \\
\text{s.t.} & 4x_1 + x_2 + 3x_3 \leqslant 30, \\
& x_1 + 2x_2 + 3x_3 \leqslant 40, \\
& x_1, x_2, x_3 \geqslant 0.
\end{cases}
$$

解 上述线性规划问题的对偶问题为

$$
\begin{cases}
\min & 30y_1 + 40y_2; \\
\text{s.t.} & 4y_1 + y_2 \geqslant 4, \\
& y_1 + 2y_2 \geqslant 3, \\
& y_1 + y_2 \geqslant 2, \\
& y_1, y_2 \geqslant 0.
\end{cases}
$$

我们用图解法求解此对偶问题,如图 4.3.1 所示,得知其最优解为 $\bar{y}=(1,1)^{\mathrm{T}}$.

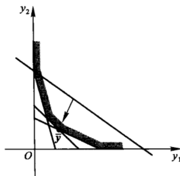


图 4.3.1　例 4.3.3 中问题的对偶问题的图解法

注意到 $\bar{y}_1=1>0$, $\bar{y}_2=1>0$, $4\bar{y}_1+\bar{y}_2>4$,根据互补松弛条件,原问题的最优解 $\bar{x}=(\bar{x}_1,\bar{x}_2,\bar{x}_3)^{\mathrm{T}}$ 应满足

$$\begin{cases}4\bar{x}_1+\bar{x}_2+3\bar{x}_3=30,\\ \bar{x}_1+2\bar{x}_2+3\bar{x}_3=40,\\ \bar{x}_1=0.\end{cases}$$

由此解得原问题的最优解 $\bar{x}=(0,10,20/3)^{\mathrm{T}}$,最优值为 70. □

最后,利用实例营养问题来解释对偶线性规划的经济意义. 设 $\bar{x}$ 和 $\bar{y}$ 分别为问题(4.3.3)和问题(4.3.4)的最优解,则

(1) $\bar{y}$ 的第 i 个分量 $\bar{y}_i$ 称为第 i 种营养成分的影子价格(shadow price),此价格表示饲养场从饲料公司得到一个单位的第 i 种营养成分所要支付的最大的价钱;

(2) 当第 i 种营养成分过剩(即 $\sum\limits_{j=1}^{n}a_{ij}\bar{x}_j>b_i$)时,互补松弛条件要求 $\bar{y}_i=0$,它表示饲养场对第 i 种营养成分不愿支持任何费用;

(3) 当第 j 种配料的价格过高(即 $\sum\limits_{i=1}^{m}a_{ij}\bar{y}_j<c_i$)时,互补松弛条件要求 $\bar{x}_j=0$,这表示饲养场不愿再使用第 j 种配料.

4.4　对偶单纯形法

设 $\bar{x}$ 为(LP)中关于基 $\boldsymbol{B}$ 的基本解,令 $\bar{y}=(\boldsymbol{c}_B^{\mathrm{T}}\boldsymbol{B}^{-1})^{\mathrm{T}}$. 若 $\bar{x}$ 和 $\bar{y}$ 分别是(LP)和(DP)的可行解,则 $\boldsymbol{A}^{\mathrm{T}}\bar{y}\leqslant\boldsymbol{c}$,这等价于

$$\boldsymbol{\pi} = (\boldsymbol{c}_B^{\mathrm{T}}\boldsymbol{B}^{-1}\boldsymbol{A} - \boldsymbol{c}^{\mathrm{T}})^{\mathrm{T}} = \boldsymbol{A}^{\mathrm{T}}\bar{\boldsymbol{y}} - \boldsymbol{c} \leqslant \boldsymbol{0},$$

即 $\bar{\boldsymbol{x}}$ 对应的检验数全部非正,故由定理 4.2.1 知,$\bar{\boldsymbol{x}}$ 是(LP)的最优解. 而

$$\boldsymbol{b}^{\mathrm{T}}\bar{\boldsymbol{y}} = \bar{\boldsymbol{y}}^{\mathrm{T}}\boldsymbol{b} = \boldsymbol{c}_B^{\mathrm{T}}\boldsymbol{B}^{-1}\boldsymbol{b} = \boldsymbol{c}^{\mathrm{T}}\bar{\boldsymbol{x}},$$

所以由定理 4.3.5 知,$\bar{\boldsymbol{y}}$ 是(DP)的最优解. 这不但说明可以由(LP)的最优解得出(DP)的最优解,而且表明:(LP)中关于基 $\boldsymbol{B}$ 的基本解 $\bar{\boldsymbol{x}}$ 对应的检验数全部非正当且仅当$(\boldsymbol{c}_B^{\mathrm{T}}\boldsymbol{B}^{-1})^{\mathrm{T}}$ 为(DP)的可行解. 因此,我们把(LP)中检验数全部非正的基本解称为对偶可行解(dual feasible solution)或正则解(regular solution),相应的基称为对偶可行基(dual feasible basis)或正则基(regular basis).

于是可知,单纯形法其实是从一个基本解迭代到另一个基本解,在迭代过程中始终保持可行性,使其对偶不可行性(非正则性)逐步消失,一直到对偶可行性(正则性)被满足,便是最优解.

因为(LP)和(DP)互为对偶问题,所以基于对称的想法,我们可以给出求解(LP)的这样一个迭代过程:从一个基本解迭代到另一个基本解,在迭代过程中保持对偶可行性(正则性),使其不可行性逐步消失,一旦满足可行性便是最优解. 这就是对偶单纯形法(dual simplex method)的基本思想. 相对而言,把前面介绍的单纯形法称为原始单纯形法(prime simplex method).

设已知(LP)的一个正则基 $\boldsymbol{B} = (\boldsymbol{p}_{j_1}, \boldsymbol{p}_{j_2}, \cdots, \boldsymbol{p}_{j_m})$,对应的正则解为 $\bar{\boldsymbol{x}}$,仍用 R 表示非基变量的指标集,即

$$R = \{1, 2, \cdots, n\} \setminus \{j_1, j_2, \cdots, j_m\}.$$

设(LP)关于基 $\boldsymbol{B}$ 的典式为

$$\begin{cases} \min & f = \bar{f} - \sum\limits_{j \in R} \pi_j x_j; \\ \text{s.t.} & x_{j_i} = b_{i0} - \sum\limits_{j \in R} b_{ij} x_j, i = 1, 2, \cdots, m, \\ & x_j \geqslant 0, j = 1, 2, \cdots, n. \end{cases} \tag{4.4.1}$$

因 $\bar{\boldsymbol{x}}$ 是正则解,故典式(4.4.1)中的检验数

$$\pi_j \leqslant 0, \forall j \in R.$$

如果 $b_{i0} \geqslant 0 (i = 1, 2, \cdots, m)$,则 $\bar{\boldsymbol{x}}$ 为(LP)的最优解. 否则,即有某些 $b_{i0} < 0$,则可选满足

$$\min\{b_{i0} \mid b_{i0} < 0\} = b_{s0} \tag{4.4.2}$$

的 x_{j_s} 为离基变量. 以下分两种情况讨论.

(1) 若所有 $b_{sj} \geqslant 0 (j \in R)$,则(LP)无可行解.

这是因为,假如(LP)有可行解 $\hat{\boldsymbol{x}} = (\hat{x}_1, \hat{x}_2, \cdots, \hat{x}_n)^{\mathrm{T}}$,则有

$$\hat{x}_{j_s} = b_{s0} - \sum_{j \in R} b_{sj} \hat{x}_j.$$

由 $b_{s0} < 0, b_{sj} \geqslant 0 (j \in R)$ 和 $\hat{x}_j \geqslant 0 (j \in R)$ 知 $\hat{x}_{j_s} < 0$,此与 $\hat{\boldsymbol{x}}$ 为可行解矛盾.

(2) 若存在某些 $b_{sj}<0$,则可在这些元素所在列对应的非基变量中选取进基变量.

假设 x_r 为进基变量,因 $b_{sr}\neq 0$,故由定理 4.2.4 知,约束矩阵 $\boldsymbol{A}$ 的列向量组

$$\boldsymbol{p}_{j_1},\cdots,\boldsymbol{p}_{j_{s-1}},\boldsymbol{p}_r,\boldsymbol{p}_{j_{s+1}},\cdots,\boldsymbol{p}_{j_m}$$

线性无关,即上述列向量构成新基 $\widetilde{\boldsymbol{B}}$. 将典式(4.4.1)作 $\{s,r\}$ 旋转变换:先解 x_r,得

$$x_r=\frac{b_{s0}}{b_{sr}}-\sum_{j\in R\setminus\{r\}}\frac{b_{sj}}{b_{sr}}x_j-\frac{1}{b_{sr}}x_{j_s},$$

再回代到典式(4.4.1)中,得

$$x_{j_i}=\left(b_{i0}-\frac{b_{s0}}{b_{sr}}b_{ir}\right)-\sum_{j\in R\setminus\{r\}}\left(b_{ij}-\frac{b_{sj}}{b_{sr}}b_{ir}\right)x_j+\frac{b_{ir}}{b_{sr}}x_{j_s},i\neq s,$$

$$f=\left(\bar{f}-\pi_r\frac{b_{s0}}{b_{sr}}\right)-\sum_{j\in R\setminus\{r\}}\left(\pi_j-\pi_r\frac{b_{sj}}{b_{sr}}\right)x_j-\left(-\frac{\pi_r}{b_{sr}}\right)x_{j_s}. \tag{4.4.3}$$

为使新基 $\widetilde{\boldsymbol{B}}$ 成为正则基,则应使(4.4.3)式中新的检验数非正,即

$$-\frac{\pi_r}{b_{sr}}\leqslant 0, \tag{4.4.4}$$

$$\pi_j-\pi_r\frac{b_{sj}}{b_{sr}}\leqslant 0,\ \forall j\in R\setminus\{r\}. \tag{4.4.5}$$

要使(4.4.4)式成立,因 $\pi_r\leqslant 0$,故要求 $b_{sr}<0$. 从而,当 $b_{sj}\geqslant 0$ 时,(4.4.5)式自然成立;当 $b_{sj}<0$ 时,要使(4.4.5)成立,则要求

$$\frac{\pi_r}{b_{sr}}\leqslant\frac{\pi_j}{b_{sj}},$$

由此可知,只要选择 x_r 满足

$$\frac{\pi_r}{b_{sr}}=\min\left\{\frac{\pi_j}{b_{sj}}\,\middle|\,b_{sj}<0,j\in R\right\}, \tag{4.4.6}$$

那么(4.4.4)式和(4.4.5)式均能被满足. 所以,只需分别按(4.4.2)式和(4.4.6)式选取离基变量 x_{j_s} 和进基变量 x_r,就能得到关于基 $\widetilde{\boldsymbol{B}}$ 的正则解 $\tilde{\boldsymbol{x}}$.

注意到 $\pi_r\leqslant 0,b_{s0}<0,b_{sr}<0$,从而由(4.4.3)式知

$$\boldsymbol{c}^{\mathrm{T}}\tilde{\boldsymbol{x}}=\bar{f}-\pi_r\frac{b_{s0}}{b_{sr}}\geqslant\bar{f}=\boldsymbol{c}^{\mathrm{T}}\bar{\boldsymbol{x}}.$$

如果 $\pi_r<0$,则必有 $\boldsymbol{c}^{\mathrm{T}}\tilde{\boldsymbol{x}}>\boldsymbol{c}^{\mathrm{T}}\bar{\boldsymbol{x}}$. 因此,在对偶单纯形迭代中,目标函数值是单调增加的,这与原始单纯形迭代的情况正好相反,其原因在于,对偶单纯形迭代是从(LP)的可行域外部逐步趋于最优解. 并且,如果(LP)的每一个正则解中非基变量对应的检验数都是负的(此时称(DP)是非退化的),那么每迭代一次,目标函数值严格单调增加,从而出现过的正则解不会再次出现,于是经过有限次迭代

后，或者得到(LP)的最优解，或者判定(LP)无可行解.

以上所述的从一个正则解迭代到另一个正则解的方法，与原始单纯形法中从一个基本可行解迭代到另一个基本可行解的方法是类似的，都是作$\{s,r\}$旋转变换，所不同的是离基变量和进基变量的选取方式. 因此，对偶单纯形法也可以在单纯形表上进行.

算法 4-2(对偶单纯形法)

Step 1　选取(LP)的一个关于正则基 $\boldsymbol{B}$ 的正则解 $\bar{\boldsymbol{x}}$，列出单纯形表 $T(\boldsymbol{B})$.

Step 2　若 $b_{i0}\geqslant 0(i=1,2,\cdots,m)$，则 $\bar{\boldsymbol{x}}$ 是最优解，算法结束；否则，按

$$j_s = \min\{j_k \mid b_{k0} = \min_{b_{i0}<0} b_{i0}\}$$

选取 x_{j_s}为离基变量.

Step 3　若 $b_{sj}\geqslant 0(j\in R)$，则(LP)无可行解，算法终止；否则，按

$$r = \min\left\{l \,\middle|\, \frac{\pi_l}{b_{sl}} = \min_{b_{sj}<0}\frac{\pi_j}{b_{sj}}\right\}$$

选取 x_r 为进基变量.

Step 4　以 b_{sr}为主元进行$\{s,r\}$旋转变换，得到新的单纯形表 $T(\widetilde{\boldsymbol{B}})$，以 $\widetilde{\boldsymbol{B}}$ 取代 $\boldsymbol{B}$，返回 Step 2.

例 4.4.1　用对偶单纯形法求解线性规划问题

$$\begin{cases}\min & f = x_1 + 3x_2 + x_3;\\ \text{s.t.} & 2x_1 + x_2 + x_3 \geqslant 3,\\ & 3x_1 + 2x_2 \geqslant 4,\\ & x_1 + 2x_2 - x_3 \geqslant 1,\\ & x_1, x_2, x_3 \geqslant 0.\end{cases}$$

解　引进变量 x_4,x_5,x_6，将给定的线性规划问题化为标准形式：

$$\begin{cases}\min & f = x_1 + 3x_2 + x_3;\\ \text{s.t.} & -2x_1 - x_2 - x_3 + x_4 = -3,\\ & -3x_1 - 2x_2 + x_5 = -4,\\ & -x_1 - 2x_2 + x_3 + x_6 = -1,\\ & x_1,x_2,x_3,x_4,x_5,x_6 \geqslant 0.\end{cases}$$

$\boldsymbol{B}=(\boldsymbol{p}_4,\boldsymbol{p}_5,\boldsymbol{p}_6)$ 是一个明显的正则基，对应的初始正则解 $\bar{\boldsymbol{x}}=(0,0,0,-3,-4,-1)^{\mathrm{T}}$. 相应的单纯形表为表 4.4.1.

表 4.4.1 例 4.4.1 的初始单纯形表

	x_1	x_2	x_3	x_4	x_5	x_6	
x_4	-2	-1	-1	1	0	0	-3
x_5	-3^*	-2	0	0	1	0	-4
x_6	-1	-2	1	0	0	1	-1
f	-1	-3	-1	0	0	0	0

因为 $b_{10} = -3 < 0, b_{20} = -4 < 0, b_{30} = -3 < 0$,所以取 x_5 为离基变量. 而

$$\min\left\{\frac{\pi_j}{b_{2j}}\middle| b_{2j} < 0\right\} = \min\left\{\frac{-1}{-3}, \frac{-3}{-2}\right\} = \frac{1}{3} = \frac{\pi_1}{b_{21}},$$

故取 x_1 为进基变量. 以 $b_{21} = -3$ 为主元作{2,1}旋转变换,得表 4.4.2.

表 4.4.2 例 4.4.1 迭代一次后的单纯形表

	x_1	x_2	x_3	x_4	x_5	x_6	
x_4	0	1/3	-1	1	$-2/3^*$	0	$-1/3$
x_1	1	2/3	0	0	$-1/3$	0	4/3
x_6	0	$-4/3$	1	0	$-1/3$	1	1/3
f	0	$-7/3$	-1	0	$-1/3$	0	4/3

同理,再在表 4.4.2 中进行{1,5}旋转变换,得表 4.4.3. 而表 4.4.3 对应的正则解是可行解,即知给定的问题的最优解为 $\boldsymbol{x} = \left(\frac{3}{2},0,0\right)^{\mathrm{T}}$,最优值为$\frac{3}{2}$.

表 4.4.3 例 4.4.1 迭代两次后的最优单纯形表

	x_1	x_2	x_3	x_4	x_5	x_6	
x_5	0	$-1/2$	3/2	$-3/2$	1	0	1/2
x_1	1	1/2	1/2	$-1/2$	0	0	3/2
x_6	0	$-3/2$	3/2	$-1/2$	0	1	1/2
f	0	$-5/2$	$-1/2$	$-1/2$	0	0	3/2

□

如果对偶问题(DP)是退化的,那么用对偶单纯形法求解(LP)就可能导致基的循环. 为了避免发生这种现象,我们可以采用下面的规则选取离基变量和进

基变量.

（1）离基规则：由

$$j_s = \min\{j_i \mid b_{i0} < 0\}$$

确定 x_{j_s} 为离基变量；

（2）进基规则不变：即仍由

$$r = \min\left\{l \,\middle|\, \frac{\pi_l}{b_{sl}} = \min_{b_{sj}<0} \frac{\pi_j}{b_{sj}}\right\}$$

选取 x_r 为进基变量.

这里的离基规则和进基规则也称为 Bland 规则.

最后，我们讨论求初始正则解的方法.

我们知道，用对偶单纯形法求解（LP）时，必须先知道一个初始正则解. 如果所给的（LP）不含有明显的正则基，如何进行对偶单纯形迭代呢？下面来回答这个问题.

设已知（LP）的关于基 $\boldsymbol{B}$ 的基本解 $\bar{\boldsymbol{x}}$，它不是正则解（也不必是可行解），对应的典式为（4.4.1）式，现在引进一个人工约束（artificial constraint）

$$\sum_{j\in R} x_j + x_{n+1} = M,$$

其中 M 表示充分大的正数，x_{n+1} 为引进的变量. 把这个约束添加到（4.4.1）式中得到一个新的线性规划问题

$$\begin{cases} \min & f = \bar{f} - \sum\limits_{j\in R} \pi_j x_j; \\ \text{s.t.} & x_{j_i} = b_{i0} - \sum\limits_{j\in R} b_{ij} x_j,\ i = 1,2,\cdots,m, \\ & x_{n+1} = M - \sum\limits_{j\in R} x_j, \\ & x_j \geqslant 0,\ j = 1,2,\cdots,n,n+1. \end{cases} \tag{4.4.7}$$

问题（4.4.7）称为（LP）关于基 $\boldsymbol{B}$ 的扩充问题（extended problem）.

对于扩充问题（4.4.7），按下述方法处理，即可得出它的一个正则解. 设

$$\pi_r = \max\{\pi_j \mid \pi_j > 0, j \in R\},$$

选取 x_r 为进基变量，并指定 x_{n+1} 为离基变量，以 $b_{m+1,r}$ 为主元作 $\{m+1, r\}$ 旋转变换，得到新的典式. 易知新的典式中目标函数的表达为

$$f = (\bar{f} - \pi_r M) - \sum_{j\in R\setminus\{r\}} (\pi_j - \pi_r) x_j - (-\pi_r) x_{n+1},$$

其中检验数

$$\pi_j - \pi_r \leqslant 0,\ \forall j \in R\setminus\{r\};\ -\pi_r < 0.$$

所以，经上述迭代所得的新的基本解便是（4.4.7）式的正则解.

扩充问题（4.4.7）有了初始正则解后，便可开始对偶单纯形迭代. 迭代结果

有且仅有下列三种可能情形：

（1）扩充问题无可行解，则(LP)也无可行解；

（2）扩充问题有最优解$(\tilde{x}_1,\tilde{x}_2,\cdots,\tilde{x}_n,\tilde{x}_{n+1})^{\mathrm{T}}$，且扩充问题的最优值与$M$无关，则$(\tilde{x}_1,\tilde{x}_2,\cdots,\tilde{x}_n)^{\mathrm{T}}$是(LP)的最优解；

（3）扩充问题有最优解$(\tilde{x}_1,\tilde{x}_2,\cdots,\tilde{x}_n,\tilde{x}_{n+1})^{\mathrm{T}}$，但扩充问题的最优解值与$M$有关，则(LP)无最优解.

以上三种可能情形的证明留作习题.

例 4.4.2 求解线性规划问题

$$\begin{cases}\min & f=-2x_2+x_3;\\ \text{s.t.} & x_1+x_2+x_3=5,\\ & x_2-x_3+x_4=-1,\\ & x_1,x_2,x_3,x_4\geqslant 0.\end{cases}$$

解 显然$(\boldsymbol{p}_1,\boldsymbol{p}_4)$为一个基，但既不是可行基也不是正则基，因为对应的基本解$\bar{\boldsymbol{x}}=(5,0,0,-1)^{\mathrm{T}}$，且检验数$\pi_2=2>0$. 增加人工约束

$$x_2+x_3+x_5=M,$$

列出扩充问题的初始单纯形表 4.4.4.

表 4.4.4 例 4.4.2 扩充问题的初始单纯形表

	x_1	x_2	x_3	x_4	x_5	
x_1	1	1	2	0	0	5
x_4	0	1	-1	1	0	-1
x_5	0	1^*	1	0	1	M
f	0	2	-3	0	0	0

首次迭代的离基变量指定为x_5，并由

$$\max\{\pi_j\mid\pi_j>0\}=2=\pi_2$$

确定x_2为进基变量，进行{3,2}旋转变换，得表 4.4.5.

从单纯形表 4.4.5 开始即按对偶单纯形法进行迭代. 因为$b_{10}=-M+5<0$，$b_{20}=-M-1<0$，所以选x_4为离基变量，而

$$\min\left\{\frac{\pi_j}{b_{2j}}\middle|b_{2j}<0\right\}=\min\left\{\frac{-5}{-2},\frac{-2}{-1}\right\}=2=\frac{\pi_5}{b_{25}},$$

故取x_5为进基变量. 作{2,5}旋转变换，得表 4.4.6.

表 4.4.5　例 4.4.2 扩充问题迭代一次的单纯形表

	x_1	x_2	x_3	x_4	x_5	
x_1	1	0	1	0	-1	$-M+5$
x_4	0	0	-2	1	-1^*	$-M-1$
x_2	0	1	1	0	1	M
f	0	0	-5	0	-2	$-2M$

表 4.4.6　例 4.4.2 扩充问题迭代两次的单纯形表

	x_1	x_2	x_3	x_4	x_5	
x_1	1	0	3	-1	0	6
x_5	0	0	2	-1	1	$M+1$
x_2	0	1	-1^*	1	0	-1
f	0	0	-1	-2	0	2

同理,在表 4.4.6 中再进行$\{3,3\}$旋转变换,得表 4.4.7.

表 4.4.7　例 4.4.2 扩充问题迭代三次后的最优单纯形表

	x_1	x_2	x_3	x_4	x_5	
x_1	1	3	0	2	0	3
x_5	0	2	0	1	1	$M-1$
x_3	0	-1	1	-1	0	1
f	0	-1	0	-3	0	3

单纯形表 4.4.7 对应的正则解已是可行解,从而是扩充问题的最优解,且对应的目标函数值与 M 无关,因此得到给定的问题的最优解

$$\bar{x} = (3,0,1,0)^{\mathrm{T}},$$

最优值为 3. □

习　题　四

1. 将下列线性规划问题化为标准形式:

(1) $\begin{cases} \max & -2x_1+x_2-2x_3; \\ \text{s.t.} & -x_1+x_2+x_3=4, \\ & -x_1+x_2-x_3\leqslant 6, \\ & x_1\leqslant 0, x_2\geqslant 0. \end{cases}$

(2) $\min \sum_{i=1}^{m} \left| \sum_{j=1}^{n} a_{ij}x_j - b_i \right|$.

2. 用图解法求解下列线性规划问题:

(1) $\begin{cases} \max & 2x_1+7x_2; \\ \text{s.t.} & x_1+2x_2\leqslant 16, \\ & 2x_1+x_2\leqslant 12, \\ & x_1,x_2\geqslant 0. \end{cases}$

(2) $\begin{cases} \min & 2x_1+6x_2; \\ \text{s.t.} & x_1+3x_2\geqslant 3, \\ & 2x_1-x_2\geqslant 2, \\ & x_1,x_2\geqslant 0. \end{cases}$

(3) $\begin{cases} \min & 2x_1+6x_2+8x_3; \\ \text{s.t.} & x_1+2x_2+x_3=5, \\ & 2x_1+3x_2+x_3=6, \\ & x_1,x_2,x_3\geqslant 0. \end{cases}$

3. 对于线性规划问题

$$\begin{cases} \min & 4x_1+5x_2; \\ \text{s.t.} & x_1+x_2+x_3=1, \\ & 2x_1+3x_2=1, \\ & x_1,x_2,x_3\geqslant 0, \end{cases}$$

求出全部基本可行解,并比较出最优解.

4. 考虑线性规划问题

$$\begin{cases} \min & x_1+\beta x_2; \\ \text{s.t.} & -x_1+x_2\leqslant 1, \\ & -x_1+2x_2\leqslant 4, \\ & x_1,x_2\geqslant 0, \end{cases}$$

试用图解法讨论,当 β 取何值时,上述问题

(1) 有惟一的最优解;

(2) 有无穷多个最优解;

(3) 无最优解.

5. 设 $\boldsymbol{x}_0$ 是(LP)中关于不同基 $\boldsymbol{B}_1,\boldsymbol{B}_2$ 的基本可行解,证明:$\boldsymbol{x}_0$ 为(LP)的退化的基本可行解.

6. 设 $\boldsymbol{A}\in\mathbb{R}^{m\times n}$ 是行满秩矩阵,$b\in\mathbb{R}^n$,且 $\bar{\boldsymbol{x}}$ 是多胞形

$$K=\{\boldsymbol{x}\mid \boldsymbol{Ax}=\boldsymbol{b},\boldsymbol{x}\geqslant\boldsymbol{0}\}$$

的一个极点. 证明:存在向量 $\boldsymbol{c}\in\mathbb{R}^n$,使得 $\bar{\boldsymbol{x}}$ 是线性规划问题

$$\begin{cases} \min & \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}; \\ \text{s.t.} & \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}, \\ & \boldsymbol{x} \geqslant \boldsymbol{0} \end{cases}$$

的惟一的最优解.

7. 证明:(LP)的非退化基本可行解 $\bar{\boldsymbol{x}}$ 是其惟一最优解的充要条件是 $\bar{\boldsymbol{x}}$ 的所有非基变量对应的检验数均为负数.

8. 用单纯形法求解下列线性规划问题:

(1) $$\begin{cases} \max & x_1 + x_2 + x_3 + x_4; \\ \text{s.t.} & x_1 + x_2 \leqslant 2, \\ & x_3 + x_4 \leqslant 5, \\ & x_1, x_2, x_3, x_4 \geqslant 0. \end{cases}$$

(2) $$\begin{cases} \min & x_1 - 2x_2 + x_4; \\ \text{s.t.} & x_1 + 3x_3 + 2x_5 = 12, \\ & x_2 - 2x_3 + x_4 = 2, \\ & x_2 + x_3 + x_5 = 5, \\ & x_1, x_2, x_3, x_4, x_5 \geqslant 0. \end{cases}$$

(3) $$\begin{cases} \min & 2x_3 - 3x_5; \\ \text{s.t.} & x_1 + x_3 - x_5 = 1, \\ & x_2 - 2x_3 + x_5 = 1, \\ & 3x_3 + x_4 - 2x_5 = 2, \\ & x_1, x_2, x_3, x_4, x_5 \geqslant 0. \end{cases}$$

(4) $$\begin{cases} \max & x_1 + x_2 + x_3; \\ \text{s.t.} & -x_1 - 2x_3 \leqslant 5, \\ & 2x_1 - 3x_2 + x_3 \leqslant 3, \\ & 2x_1 - 5x_2 + 6x_3 \leqslant 5, \\ & x_1, x_2, x_3 \geqslant 0. \end{cases}$$

9. 用 Dantzig 规则和 Bland 规则分别求解 Beale 例子.

10. 试用两阶段法求解下列线性规划问题:

(1) $$\begin{cases} \min & x_1 + x_2 + 2x_3 + 5x_4; \\ \text{s.t.} & x_1 - 2x_3 + x_4 = 1, \\ & x_1 + x_2 + x_3 = 4, \\ & 2x_1 - 4x_3 + 2x_4 = 2, \\ & x_1, x_2, x_3, x_4 \geqslant 0. \end{cases}$$

(2) $\begin{cases} \min & x_1 + 2x_2 + 3x_3; \\ \text{s.t.} & x_1 - 2x_2 + 4x_3 = 4, \\ & 4x_1 - 9x_2 - 3x_3 = 16, \\ & x_1, x_2, x_3 \geqslant 0. \end{cases}$

(3) $\begin{cases} \max & 10x_1 + 15x_2 + 12x_3; \\ \text{s.t.} & 5x_1 + 3x_2 + x_3 \leqslant 9, \\ & 2x_1 + x_2 + x_3 \geqslant 5, \\ & -5x_1 + 6x_2 + 15x_3 \leqslant 15, \\ & x_1, x_2, x_3 \geqslant 0. \end{cases}$

11. 写出下列线性规划问题的对偶问题:

(1) $\begin{cases} \min & 2x_1 + x_2 + 3x_3 + x_4; \\ \text{s.t.} & x_1 + x_2 + x_3 + x_4 \leqslant 5, \\ & 2x_1 - x_2 + 3x_3 = -4, \\ & x_1 - x_3 + x_4 \geqslant 1, \\ & x_1 \geqslant 0, x_3 \geqslant 0. \end{cases}$

(2) $\begin{cases} \min & 3x_1 + 2x_2 - 3x_3 + 4x_4; \\ \text{s.t.} & x_1 - 2x_2 + 3x_3 + 4x_4 \leqslant 3, \\ & x_2 + 3x_3 + 4x_4 \geqslant -5, \\ & 2x_1 - 3x_2 - 7x_3 - 4x_4 = 2, \\ & x_1 \geqslant 0, x_4 \leqslant 0. \end{cases}$

12. 写出下列线性规划问题的对偶问题及其互补松弛条件:

(1) $\begin{cases} \min & \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}; \\ \text{s.t.} & \boldsymbol{A}\boldsymbol{x} = \boldsymbol{0}, \\ & \boldsymbol{e}^{\mathrm{T}}\boldsymbol{x} = 1, \\ & \boldsymbol{x} \geqslant \boldsymbol{0}. \end{cases}$ (2) $\begin{cases} \min & \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}; \\ \text{s.t.} & \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}, \\ & \boldsymbol{l} \leqslant \boldsymbol{x} \leqslant \boldsymbol{u}. \end{cases}$

13. 写出下面运输问题的对偶问题:

$$\begin{cases} \min & \sum_{i=1}^{m} \sum_{j=1}^{n} c_{ij} x_{ij}; \\ \text{s.t.} & \sum_{j=1}^{n} x_{ij} = a_i, i = 1,2,\cdots,m, \\ & \sum_{i=1}^{m} x_{ij} = b_j, j = 1,2,\cdots,n, \\ & x_{ij} \geqslant 0, i = 1,2,\cdots,m; j = 1,2,\cdots,n. \end{cases}$$

14. 假设线性规划问题(LP)有最优解,证明:对一切 $\boldsymbol{d} \in \mathbb{R}^m$,线性规划问题

$$\begin{cases} \min & \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}; \\ \text{s.t.} & \boldsymbol{A}\boldsymbol{x} = \boldsymbol{d}, \\ & \boldsymbol{x} \geqslant \boldsymbol{0} \end{cases}$$

不会出现有可行解而无最优解的情况.

15. 对于线性规划问题(LP),假设 $\boldsymbol{A}^{\mathrm{T}} = \boldsymbol{A}, \boldsymbol{c} = \boldsymbol{b}$,试证:若 $\boldsymbol{x}_0$ 为(LP)的可行解,则 $\boldsymbol{x}_0$ 必是最优解.

16. 对于互为对偶问题的问题(4.3.1)和(4.3.2),记

$$\varphi(\boldsymbol{x},\boldsymbol{y}) = \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x} - \boldsymbol{y}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x} + \boldsymbol{b}^{\mathrm{T}}\boldsymbol{y},$$

设 $\bar{\boldsymbol{x}} \geqslant \boldsymbol{0}$, $\bar{\boldsymbol{y}} \geqslant \boldsymbol{0}$,证明:$\bar{\boldsymbol{x}}$ 和 $\bar{\boldsymbol{y}}$ 分别为问题(4.3.1)和(4.3.2)的最优解的充要条件是$(\bar{\boldsymbol{x}}, \bar{\boldsymbol{y}})$为 $\varphi(\boldsymbol{x},\boldsymbol{y})$ 的鞍点,即对一切 $\boldsymbol{x} \in \mathbb{R}^n, \boldsymbol{y} \in \mathbb{R}^m, \boldsymbol{x} \geqslant \boldsymbol{0}, \boldsymbol{y} \geqslant \boldsymbol{0}$,有

$$\varphi(\bar{\boldsymbol{x}},\boldsymbol{y}) \leqslant \varphi(\bar{\boldsymbol{x}},\bar{\boldsymbol{y}}) \leqslant \varphi(\boldsymbol{x},\bar{\boldsymbol{y}}).$$

17. 应用对偶理论证明下面线性规划问题无最优解:

$$\begin{cases} \min & x_1 - x_2 + x_3; \\ \text{s.t.} & x_1 - x_3 \geqslant 4, \\ & x_1 - x_2 + 2x_3 \geqslant 4, \\ & x_1, x_2, x_3 \geqslant 0. \end{cases}$$

18. 应用对偶理论证明线性规划问题

$$\begin{cases} \max & 4x_1 + 5x_2 + 9x_3; \\ \text{s.t.} & x_1 + x_2 + 2x_3 \leqslant 16, \\ & 7x_1 + 5x_2 + 3x_3 \leqslant 25, \\ & x_1, x_2, x_3 \geqslant 0 \end{cases}$$

及其对偶问题都有最优解,并求最优解的上界和下界.

19. 已知线性规划问题

$$\begin{cases} \max & x_1 + 2x_2 + 3x_3 + 4x_4; \\ \text{s.t.} & x_1 + 2x_2 + 2x_3 + 3x_4 \leqslant 20, \\ & 2x_1 + x_2 + 3x_3 + 2x_4 \leqslant 20, \\ & x_1, x_2, x_3, x_4 \geqslant 0 \end{cases}$$

的对偶问题的最优解为 $\bar{\boldsymbol{y}} = \left(\dfrac{6}{5}, \dfrac{1}{5}\right)^{\mathrm{T}}$,试用互补松弛条件求原问题的最优解.

20. 已知线性规划问题

$$\begin{cases} \min & 5x_1 + 21x_3; \\ \text{s.t.} & x_1 - x_2 + 6x_3 \geqslant \alpha, \\ & x_1 + x_2 + 2x_3 \geqslant 1, \\ & x_1, x_2, x_3 \geqslant 0 \end{cases}$$

的最优解为 $\bar{\boldsymbol{x}} = \left(\dfrac{1}{2}, 0, \dfrac{1}{4}\right)^{\mathrm{T}}$,求对偶问题的最优解和参数 α 的值.

21. 把线性问题(4.3.1)化成标准形式

$$\begin{cases} \min & \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}; \\ \text{s.t.} & \boldsymbol{A}\boldsymbol{x} - \boldsymbol{z} = \boldsymbol{b}, \\ & \boldsymbol{x} \geqslant \boldsymbol{0}, \boldsymbol{z} \geqslant \boldsymbol{0}. \end{cases}$$

再用单纯形法求解这一问题.证明:在最优单纯形表中,$\boldsymbol{z}$ 变量对应的检验数反号所组成的向量就是对偶问题(4.3.2)的最优解.

22. 用对偶单纯形法求解下列线性规划问题:

$$(1)\begin{cases} \min & 12x_1 + 18x_2 + 16x_3 + 12x_4; \\ \text{s.t.} & 2x_1 + x_2 + 4x_3 \geqslant 2, \\ & 2x_1 + 2x_2 + 4x_4 \geqslant 3, \\ & x_1, x_2, x_3, x_4 \geqslant 0. \end{cases}$$

$$(2)\begin{cases} \min & x_1 + 3x_2 + x_3, \\ \text{s.t.} & 2x_1 + x_2 + x_3 \geqslant 3, \\ & 3x_1 + 2x_2 \geqslant 4, \\ & x_1 + 2x_2 - x_3 \geqslant 1, \\ & x_1, x_2, x_3 \geqslant 0. \end{cases}$$

$$(3)\begin{cases} \min & 3x_1 + 2x_2 + x_3; \\ \text{s.t.} & x_1 + x_2 + x_3 \leqslant 6, \\ & x_1 - x_3 \geqslant 4, \\ & x_2 - x_3 \geqslant 3, \\ & x_1, x_2, x_3 \geqslant 0. \end{cases}$$

23. 对于(LP)关于基 $\boldsymbol{B}$ 的扩充问题(4.4.7),在有了初始正则解后,便开始对偶单纯形迭代.证明以下三个结论:

(1) 若问题(4.4.7)无可行解,则(LP)无可行解;

(2) 若问题(4.4.7)有最优解$(\bar{x}_1, \bar{x}_2, \cdots, \bar{x}_n, \bar{x}_{n+1})^{\mathrm{T}}$,且其最优解与 M 无关,则$(\bar{x}_1, \bar{x}_2, \cdots, \bar{x}_n)^{\mathrm{T}}$ 是(LP)的最优解;

(3) 若问题(4.4.7)的最优值与 M 有关,则(LP)无最优解.

24. 利用扩充问题求解下列线性规划问题:

$$(1)\begin{cases} \min & -2x_1 + x_2; \\ \text{s.t.} & x_1 + 2x_2 + 2x_3 \leqslant 6, \\ & x_1 + x_2 + x_3 \geqslant 4, \\ & x_1, x_2, x_3 \geqslant 0. \end{cases}$$

$$(2)\begin{cases} \min & x_1 - 2x_2 - 3x_3; \\ \text{s.t.} & x_1 + x_2 - 2x_3 + 3x_4 \geqslant 5, \\ & 2x_1 - x_2 + x_3 - x_4 \geqslant 4, \\ & x_1, x_2, x_3, x_4 \geqslant 0. \end{cases}$$

$$
(3)\begin{cases}
\min & -x_1+2x_2+3x_3; \\
\text{s. t.} & 5x_1-x_2+5x_3+x_4\leqslant 17, \\
 & x_1-2x_2+x_3-x_4\geqslant 22, \\
 & x_1+x_2-x_3+x_4\geqslant -33, \\
 & x_1,x_2,x_3,x_4\geqslant 0.
\end{cases}
$$

第五章　迭代算法

在开始学习求解非线性最优化问题的算法之前，我们在这一章里先介绍迭代算法的一些概念，以便对以后各章节中的具体算法有一个一般性的理解.

5.1　下降迭代算法的基本格式

集合 X 上的迭代是指：开始时，选定一个初始点 $\boldsymbol{x}_0 \in X$，并置 $k=0$，然后按某种规则 $\mathscr{A}$ 把第 k 次迭代点 $\boldsymbol{x}_k$ 映射为后继点 $\boldsymbol{x}_{k+1} \in X$，记作 $\boldsymbol{x}_{k+1} = \mathscr{A}(\boldsymbol{x}_k)$，并称之为第 $k+1$ 次迭代. 如果得到一个迭代点列 $\{\boldsymbol{x}_k\}$，那么把其中的规则 $\mathscr{A}$ 称为迭代算法. 如果关于某个函数 $\alpha(\boldsymbol{x})$，对每个 k，都有

$$\alpha(\boldsymbol{x}_{k+1}) < \alpha(\boldsymbol{x}_k),$$

则称 $\mathscr{A}$ 为下降迭代算法（descent iterative algorithm），简称为下降算法（descent algorithm）.

求解非线性最优化问题

$$\begin{cases} \min & f(\boldsymbol{x}); \\ \text{s.t.} & \boldsymbol{x} \in S, \end{cases} \tag{5.1.1}$$

其中 $S \subseteq \mathbb{R}^n$，$f: S \to \mathbb{R}$，一般都采用下降迭代算法，这里 $\alpha(\boldsymbol{x})$ 通常取作目标函数 $f(\boldsymbol{x})$ 本身.

设 $\boldsymbol{x}_k \in \mathbb{R}^n$ 是某迭代算法的第 k 次迭代点，$\boldsymbol{x}_{k+1}$ 是第 $k+1$ 次迭代点，记

$$\Delta \boldsymbol{x}_k = \boldsymbol{x}_{k+1} - \boldsymbol{x}_k, \tag{5.1.2}$$

则有

$$\boldsymbol{x}_{k+1} = \boldsymbol{x}_k + \Delta \boldsymbol{x}_k.$$

由（5.1.2）式知，$\Delta \boldsymbol{x}_k$ 是一个以 $\boldsymbol{x}_k$ 为起点 $\boldsymbol{x}_{k+1}$ 为终点的 n 维向量. 今设 $\boldsymbol{d}_k \in \mathbb{R}^n$ 是与 $\Delta \boldsymbol{x}_k$ 同方向的向量，则必有某个 $\lambda_k \geqslant 0$，使 $\Delta \boldsymbol{x}_k = \lambda_k \boldsymbol{d}_k$，于是有

$$\boldsymbol{x}_{k+1} = \boldsymbol{x}_k + \lambda_k \boldsymbol{d}_k. \tag{5.1.3}$$

这就是求解非线性最优化问题（5.1.1）的基本迭代格式.

通常，我们把基本迭代格式（5.1.3）中的 $\boldsymbol{d}_k$ 叫做第 $k+1$ 次搜索方向（search direction），它是一个其方向从 $\boldsymbol{x}_k$ 指向 $\boldsymbol{x}_{k+1}$ 的向量；λ_k 叫做第 $k+1$ 次步长因子或沿 $\boldsymbol{d}_k$ 方向的步长（step size）. 从（5.1.3）式可以看出，使用迭代算

法求解非线性最优化问题的关键在于如何构造每一次的搜索方向和确定适当的步长.

对于无约束最优化问题(此时 $S=\mathbb{R}^n$),搜索方向通常取目标函数 f 的下降方向. 但对于约束最优化问题,迭代一般在可行域内进行,搜索方向取可行下降方向.

现在,我们给出用基本迭代格式(5.1.3)求解非线性最优化问题的一般步骤.

Step 1　选取初始数据. 选取初始点 $\boldsymbol{x}_0$,令 $k=0$.

Step 2　构造搜索方向,依照一定规则,构造 f 在点 $\boldsymbol{x}_k$ 处的下降方向(对于无约束最优化问题),或构造 f 在点 $\boldsymbol{x}_k$ 处关于 S 的可行下降方向(对于约束最优化问题)作为搜索方向 $\boldsymbol{d}_k$.

Step 3　确定搜索步长. 确定以 $\boldsymbol{x}_k$ 为起点沿搜索方向 $\boldsymbol{d}_k$ 的适当步长 λ_k,使目标函数值有某种意义的下降,通常是使

$$f(\boldsymbol{x}_k+\lambda_k\boldsymbol{d}_k)<f(\boldsymbol{x}_k).$$

Step 4　求出新迭代点. 令

$$\boldsymbol{x}_{k+1}=\boldsymbol{x}_k+\lambda_k\boldsymbol{d}_k.$$

Step 5　检验终止条件. 判定 $\boldsymbol{x}_{k+1}$ 是否满足终止条件,若满足,停止迭代,输出近似最优解 $\boldsymbol{x}_{k+1}$;否则,令 $k:=k+1$,转 Step 2.

5.2　收敛性与收敛速度

设求解非线性最优化问题(5.1.1)的一个算法 $\mathscr{A}$ 产生的迭代点列为 $\{\boldsymbol{x}_k\}$,通常要求该算法具有下面的收敛性:当 $\{\boldsymbol{x}_k\}$ 是有穷点列时,其最后一个点是问题(5.1.1)的最优解 $\bar{\boldsymbol{x}}$;当 $\{\boldsymbol{x}_k\}$ 是无穷点列时,在某种范数 $\|\cdot\|$ 的意义下它收敛于最优解 $\bar{\boldsymbol{x}}$,即有

$$\lim_{k\to\infty}\|\boldsymbol{x}_k-\bar{\boldsymbol{x}}\|=0.$$

但是,在非常多的情况下,要使算法产生的迭代点列收敛于全局最优解是比较困难的. 因此,一般把满足某些条件的点集定义为解集合,当迭代点属于这个集合时,就停止迭代. 常用的解集有以下几种:

(1) $\Omega=\{\bar{\boldsymbol{x}}\in S\mid\bar{\boldsymbol{x}}$是局部最优解或全局最优解$\}$;

(2) $\Omega=\{\bar{\boldsymbol{x}}\in S\mid\bar{\boldsymbol{x}}$是满足最优解的必要条件的点$\}$;

根据第三章介绍的最优性条件知,对无约束问题,可以定义

$$\Omega=\{\bar{\boldsymbol{x}}\in S\mid\nabla f(\bar{\boldsymbol{x}})=\boldsymbol{0}\}.$$

对约束问题,可以定义

$$\Omega = \{\bar{x} \in S \mid \bar{x}\text{ 为 K－T 点}\}.$$

(3) $\Omega = \{\bar{x} \in S \mid \bar{x}$是某意义下可接受的近似最优解$\}$,

例如,$\Omega = \{\bar{x} \in S \mid f(\bar{x}) \leqslant b\}$,其中 b 为某个可接受的目标值.

我们把以(1)和(2)中的Ω作为解集的算法,叫做具有理论收敛性的算法;把以(3)中的Ω作为解集的算法,叫做具有实用收敛性的算法.

一个算法是否收敛,常常同初始点 $\boldsymbol{x}_0$ 的选择有关.若只有当 $\boldsymbol{x}_0$ 充分接近于 $\bar{\boldsymbol{x}} \in \Omega$,由算法 $\mathscr{A}$ 产生的点列才收敛于 $\bar{\boldsymbol{x}}$,则称该算法具有局部收敛性(local convergence).若对任意的初始点 $\boldsymbol{x}_0 \in S$,由算法 $\mathscr{A}$ 产生的点列都收敛于 $\bar{\boldsymbol{x}}$,则称该算法具有全局收敛性(global convergence).

如果某算法用于求解目标函数为二次函数的无约束问题时,只需经过有限次迭代就达到最优解,则称该算法具有二次终止性(quadratic termination).由于一般的函数在最优解附近常常可以用二次函数来近似,因此具有二次终止性的算法可望在接近最优解时具有好的收敛性质.通常,具有全局收敛性或二次终止性的算法在一定意义下被认为是比较好的算法.

衡量算法好坏的另一个重要标准是收敛速度.

定义 5.2.1 设由算法 $\mathscr{A}$ 产生的迭代点列$\{\boldsymbol{x}_k\}$收敛于 $\bar{\boldsymbol{x}}$,即有

$$\lim_{k\to\infty} \| \boldsymbol{x}_k - \bar{\boldsymbol{x}} \| = 0,$$

若存在实数 $\alpha > 0$ 及一个与迭代次数 k 无关的常数 $q > 0$,使

$$\lim_{k\to\infty} \frac{\| \boldsymbol{x}_{k+1} - \bar{\boldsymbol{x}} \|}{\| \boldsymbol{x}_k - \bar{\boldsymbol{x}} \|^{\alpha}} = q,$$

则称算法 $\mathscr{A}$ 产生的迭代点列$\{\boldsymbol{x}_k\}$具有 α 阶收敛速度或称算法 $\mathscr{A}$ 是 α 阶收敛的.特别地,

(1) 当 $\alpha = 1, q > 0$ 时,迭代点列$\{\boldsymbol{x}_k\}$叫做具有线性收敛速度,或 $\mathscr{A}$ 是线性收敛的(linear convergence);

(2) 当 $1 < \alpha < 2, q > 0$ 或 $\alpha = 1, q = 0$ 时,迭代点列$\{\boldsymbol{x}_k\}$叫做具有超线性收敛速度,或 $\mathscr{A}$ 是超线性收敛的(hyperlinear convergence);

(3) 当 $\alpha = 2$ 时,迭代点列$\{\boldsymbol{x}_k\}$叫做具有二阶收敛速度,或 $\mathscr{A}$ 是二阶收敛的(quadratic convergence).

一般认为,具有超线性收敛速度和二阶收敛速度的算法是比较快速的.不过,还应该意识到,对任何一个算法,收敛性和收敛速度的理论结果并不保证算法在实际执行时一定有好的实际计算结果.一方面是由于这些理论结果本身并不能保证算法一定有好的特性,另一方面是它们忽略了计算过程中十分重要的舍入误差的影响.此外,这些理论结果通常要对函数 $f(\boldsymbol{x})$ 加上某些不易验证的限制,这些限制条件在实际中并不一定能得到满足.因此,一个最优化算法的开发还依赖于数值试验,就是说,通过对各种形式的有代表性的检验函数进行数值

计算，一个好的算法应该具有可以接受的特征. 显然，数值试验不可能以严格的数学证明保证算法具有良好的性态. 理想的情况是根据收敛性和收敛速度的理论结果来选择适当的数值试验. 图 5.2.1 给出了用于数值试验的一个著名的检验函数——Rosenbrock 函数

$$f(\boldsymbol{x}) = 100(x_2 - x_1^2)^2 + (1 - x_1)^2$$

的等值线图. 这个函数的每条等值线形成一条窄长弯曲的谷道，外形像香蕉，又称为香蕉函数，其最优点为 $\bar{\boldsymbol{x}} = (1,1)^{\mathrm{T}}$.

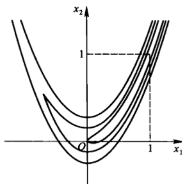


图 5.2.1　Rosenbrock 函数等值线图

下面的定理给出了算法超线性收敛的一个特征，它对于构造终止迭代所需的终止条件是有用的.

定理 5.2.1　如果点列 $\{\boldsymbol{x}_k\}$ 超线性收敛于 $\bar{\boldsymbol{x}}$，则

$$\lim_{k\to\infty} \frac{\|\boldsymbol{x}_{k+1} - \boldsymbol{x}_k\|}{\|\boldsymbol{x}_k - \bar{\boldsymbol{x}}\|} = 1.$$

证明　因 $\{\boldsymbol{x}_k\}$ 超线性收敛于 $\bar{\boldsymbol{x}}$，则有

$$\lim_{k\to\infty} \frac{\|\boldsymbol{x}_{k+1} - \bar{\boldsymbol{x}}\|}{\|\boldsymbol{x}_k - \bar{\boldsymbol{x}}\|} = 0,$$

又

$$\frac{\|\boldsymbol{x}_{k+1} - \bar{\boldsymbol{x}}\|}{\|\boldsymbol{x}_k - \bar{\boldsymbol{x}}\|} = \frac{\|(\boldsymbol{x}_{k+1} - \boldsymbol{x}_k) + (\boldsymbol{x}_k - \bar{\boldsymbol{x}})\|}{\|\boldsymbol{x}_k - \bar{\boldsymbol{x}}\|}$$

$$\geqslant \left| \frac{\|\boldsymbol{x}_{k+1} - \boldsymbol{x}_k\|}{\|\boldsymbol{x}_k - \bar{\boldsymbol{x}}\|} - \frac{\|\boldsymbol{x}_k - \bar{\boldsymbol{x}}\|}{\|\boldsymbol{x}_k - \bar{\boldsymbol{x}}\|} \right| = \left| \frac{\|\boldsymbol{x}_{k+1} - \boldsymbol{x}_k\|}{\|\boldsymbol{x}_k - \bar{\boldsymbol{x}}\|} - 1 \right|,$$

故　$\lim\limits_{k\to\infty} \dfrac{\|\boldsymbol{x}_{k+1} - \boldsymbol{x}_k\|}{\|\boldsymbol{x}_k - \bar{\boldsymbol{x}}\|} = 1.$　□

这个定理表明可以用 $\|\boldsymbol{x}_{k+1} - \boldsymbol{x}_k\|$ 来代替 $\|\boldsymbol{x}_k - \bar{\boldsymbol{x}}\|$ 给出终止判断，并且这

个估计随着 k 的增加而改善. 需要指出,该定理的逆不真.

5.3 实用终止准则

应用迭代算法时,当 $\boldsymbol{x}_k \in \Omega$ 时才终止迭代. 实践中,在许多情况下,这是一个取极限的过程,需要无限次迭代. 因此,为解决实际问题,需要规定一些实用的终止迭代过程的准则,一般称之为停步准则或终止准则(stopping criteria).

从前面的讨论可知,$\|\boldsymbol{x}_{k+1}-\boldsymbol{x}_k\|$ 是误差 $\|\boldsymbol{x}_k-\bar{\boldsymbol{x}}\|$ 的一个估计. 因此,实际中我们用 $\|\boldsymbol{x}_{k+1}-\boldsymbol{x}_k\|$ 代替 $\|\boldsymbol{x}_k-\bar{\boldsymbol{x}}\|$,这样,我们就不必预先知道最优解 $\bar{\boldsymbol{x}}$ 的信息.

现在我们给出一些常用的终止准则.

(1) 当自变量的改变量充分小时,即

$$\|\boldsymbol{x}_{k+1}-\boldsymbol{x}_k\| < \varepsilon_1 \tag{5.3.1a}$$

或

$$\frac{\|\boldsymbol{x}_{k+1}-\boldsymbol{x}_k\|}{\|\boldsymbol{x}_k\|} < \varepsilon_1 \tag{5.3.1b}$$

时,停止计算.

(2) 当函数值的下降量充分小时,即

$$|f(\boldsymbol{x}_k)-f(\boldsymbol{x}_{k+1})| < \varepsilon_1 \tag{5.3.2a}$$

或

$$\frac{|f(\boldsymbol{x}_k)-f(\boldsymbol{x}_{k+1})|}{|f(\boldsymbol{x}_k)|} < \varepsilon_1 \tag{5.3.2b}$$

时,停止计算.

终止准则(5.3.1)式和(5.3.2)式对于一些预计有较快收敛性的算法是比较理想的. 但是,在有些情况下,终止准则(5.3.1)式和(5.3.2)式是不适当的. 图 5.3.1 表明了其缺陷.

(3) 求解无约束最优化问题时,当梯度充分接近零时,即

$$\|\nabla f(\boldsymbol{x}_k)\| < \varepsilon_3 \tag{5.3.3}$$

时,停止计算.

但由于平稳点可能是鞍点,因此,在有些情况单独使用这个准则也是不适当的.

由此看来,将上述准则综合起来考虑更为合理. 1972 年,Himmelblau 提出如下终止准则,当 $\|\boldsymbol{x}_k\| \geqslant \varepsilon_2$ 和 $|f(\boldsymbol{x}_k)| \geqslant \varepsilon_2$ 时,采用

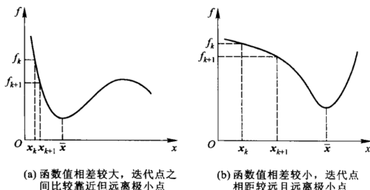


图 5.3.1　终止准则(5.3.1)和(5.3.2)不适应的例子

$$\frac{\| \boldsymbol{x}_{k+1} - \boldsymbol{x}_k \|}{\| \boldsymbol{x}_k \|} < \varepsilon_1, \frac{|f(\boldsymbol{x}_k) - f(\boldsymbol{x}_{k+1})|}{|f(\boldsymbol{x}_k)|} < \varepsilon_1, \tag{5.3.4a}$$

否则,采用

$$\| \boldsymbol{x}_{k+1} - \boldsymbol{x}_k \| < \varepsilon_1, |f(\boldsymbol{x}_k) - f(\boldsymbol{x}_{k+1})| < \varepsilon_1, \tag{5.3.4b}$$

并将(5.3.3)和(5.3.4)两式结合使用. Himmelblau 建议,一般地,可取 $\varepsilon_1 = \varepsilon_2 = 10^{-5}, \varepsilon_3 = 10^{-4}$.

最后,我们按图 5.3.2 约定算法框图元素的含义,然后在图 5.3.3 中给出 Himmelblau 终止准则(简称为 H 终止准则)的计算框图.

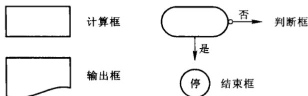


图 5.3.2　框图元素含义

框图是算法的一种直观表示,它不仅能帮助我们理解算法,而且有益于我们把算法编写成可在计算机上运行的程序. 因此,作为示范,我们还将在第六章给出一些算法的框图,使读者学习如何把一个算法的自然描述变成算法框图,以利于对算法的理解和计算机实现.

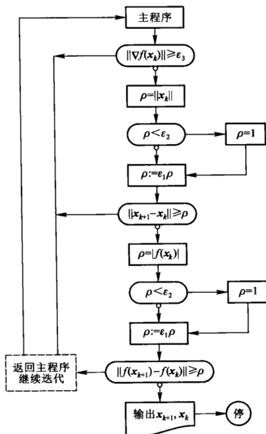


图 5.3.3　H 终止准则

习　题　五

1. 求下列各序列的收敛阶：

(1) $\left\{\frac{1}{k}\right\}$;

(2) $\left\{\left(\frac{1}{k}\right)^k\right\}$;

(3) $\{a^k\}, 0<a<1$;

(4) $\{a^{2^k}\}, |a|<1$.

2. 举例说明定理 5.2.1 的逆不真.

第六章　一维搜索

从这一章开始，我们将研究求解非线性最优化问题的具体算法．本章所讨论的一维搜索不仅是求解一维非线性最优化问题的基本算法，而且是多维非线性最优化算法的重要组成部分，它的选择是否恰当直接影响到一些算法的计算效果．

6.1　一维搜索的搜索区间

一维最优化是多维最优化的基础，求解一维最优化问题的基本思想是一维搜索，而一维搜索的重点是如何确定搜索区间．

6.1.1　一维搜索的概念

在上一章中，我们已经得到了最优化问题

$$\begin{cases}\min f(\boldsymbol{x});\\ \text{s.t. } \boldsymbol{x}\in S\end{cases} \tag{6.1.1}$$

的下降迭代算法的基本格式

$$\boldsymbol{x}_{k+1}=\boldsymbol{x}_k+\lambda_k\boldsymbol{d}_k,$$

显然，搜索方向 $\boldsymbol{d}_k$ 和步长 λ_k 构成了每一次迭代的修正量，它们是决定最优化算法好坏的重要因素．

假定给定了搜索方向 $\boldsymbol{d}_k$，从点 $\boldsymbol{x}_k$ 出发沿方向 $\boldsymbol{d}_k$ 进行搜索，要确定步长 λ_k，使

$$f(\boldsymbol{x}_{k+1})=f(\boldsymbol{x}_k+\lambda_k\boldsymbol{d}_k)<f(\boldsymbol{x}_k). \tag{6.1.2}$$

记

$$\varphi(\lambda)=f(\boldsymbol{x}_k+\lambda\boldsymbol{d}_k),$$

则(6.1.2)式等价于

$$\varphi(\lambda_k)<\varphi(0), \tag{6.1.3}$$

即确定步长 λ_k 就是单变量函数 $\varphi(\lambda)$ 的搜索问题，称为一维搜索(one dimension search)或线性搜索(linear search)．如果问题(6.1.1)是无约束最优化问题，搜索方向 $\boldsymbol{d}_k$ 取下降方向，那么，一维搜索相当于在区间 $[0,\infty)$ 上选取 $\lambda=\lambda_k$，使

(6.1.3)式成立;如果问题(6.1.1)是约束最优化问题,搜索方向 $\boldsymbol{d}_k$ 取可行下降方向,那么,一维搜索相当于在区间 $[0,\lambda_{\max}]$ 上选取 $\lambda=\lambda_k$,使(6.1.3)式成立,其中

$$\lambda_{\max} = \max\{\lambda \mid \boldsymbol{x}_k + \lambda \boldsymbol{d}_k \in S, \lambda \geqslant 0\}.$$

为方便起见,我们只讨论单变量函数

$$\varphi(\lambda) = f(\boldsymbol{x}_k + \lambda \boldsymbol{d}_k), \lambda \geqslant 0$$

的一维搜索问题. 概括地,所谓一维搜索或线性搜索是指单变量函数的最优化. 它是多变量函数最优化的基础.

通常,按对步长选取的不同原则,一维搜索分为以下两种类型.

(1) 最优一维搜索

如果求得步长 λ_k,使目标函数沿方向 $\boldsymbol{d}_k$ 达到极小,即使得

$$f(\boldsymbol{x}_k + \lambda_k \boldsymbol{d}_k) = \min_{\lambda \geqslant 0} f(\boldsymbol{x}_k + \lambda \boldsymbol{d}_k),$$

或

$$\varphi(\lambda_k) = \min_{\lambda \geqslant 0} \varphi(\lambda), \tag{6.1.4}$$

则称这样的一维搜索为最优一维搜索(optimal one dimension search),或精确一维搜索(exact one dimension search),λ_k 叫做最优步长(optimal step size). 这类一维搜索在非线性最优代方法研究中具有基本的意义,在以后几章中经常要应用它. 最优一维搜索有时也简称为一维搜索,它具有如下重要性质:

定理 6.1.1 对于问题(6.1.1),设 $f:S \to \mathbb{R}$ 是可微函数,$\boldsymbol{x}_{k+1}$ 是从 $\boldsymbol{x}_k$ 出发沿方向 $\boldsymbol{d}_k$ 作最优一维搜索得到的,则有

$$\nabla f(\boldsymbol{x}_{k+1})^{\mathrm{T}} \boldsymbol{d}_k = 0. \tag{6.1.5}$$

证明 记 $\varphi(\lambda) = f(\boldsymbol{x}_k + \lambda \boldsymbol{d}_k)$,则

$$\varphi'(\lambda) = \nabla f(\boldsymbol{x}_k + \lambda \boldsymbol{d}_k)^{\mathrm{T}} \boldsymbol{d}_k.$$

由 $\boldsymbol{x}_{k+1} = \boldsymbol{x}_k + \lambda_k \boldsymbol{d}_k$,$\lambda_k$ 满足(6.1.4)式,可知 λ_k 是 $\varphi(\lambda)$ 的极小点,因此 $\varphi'(\lambda_k)=0$,即

$$\nabla f(\boldsymbol{x}_k + \lambda_k \boldsymbol{d}_k)^{\mathrm{T}} \boldsymbol{d}_k = 0.$$

亦即(6.1.5)式成立. □

该定理指明了从点 $\boldsymbol{x}_k$ 出发沿方向 $\boldsymbol{d}_k$ 作最优一维搜索所得迭代点 $\boldsymbol{x}_{k+1}$ 的空间位置,在该点处的梯度 $\nabla f(\boldsymbol{x}_{k+1})$ 与搜索方向 $\boldsymbol{d}_k$ 正交. 图 6.1.1 给出了二维情况下,$\boldsymbol{x}_{k+1}$ 恰为该点处的等值线与搜索方向 $\boldsymbol{d}_k$ 相切的点.

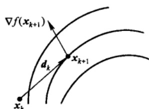


图 6.1.1 搜索方向与梯度方向正交

(2) 可接受一维搜索

如果选取 λ_k,使目标函数 f 沿方向 $\boldsymbol{d}_k$ 取得适当

的可接受的下降量，即使得下降量$f(\boldsymbol{x}_k)-f(\boldsymbol{x}_k+\lambda_k\boldsymbol{d}_k)>0$是我们可接受的，则称这样的一维搜索为可接受一维搜索(acceptable one dimension search)或非精确一维搜索(inexact one dimension search).

在实际计算中，理论上精确的最优步长一般不能求得，求近似最优步长也常需花费较大的工作量. 然而，进行可接受一维搜索往往只需很小的计算量. 因此，这类一维搜索在实用中有着重要价值.

6.1.2 确定搜索区间的进退法

考虑一维极小化问题

$$\min_{\lambda\geq 0}\varphi(\lambda). \tag{6.1.6}$$

我们引入如下搜索区间的概念

定义 6.1.1 设$\varphi:\mathbb{R}\to\mathbb{R}$，$\bar{\lambda}\in[0,\infty)$，并且

$$\varphi(\bar{\lambda})=\min_{\lambda\geq 0}\varphi(\lambda).$$

如果对于闭区间$[a,b]\subset[0,\infty)$有$\bar{\lambda}\in[a,b]$，那么称$[a,b]$是问题(6.1.6)的搜索区间.

简单地说，一个一维极小化问题的搜索区间是包含该问题最优解的一个闭区间. 通常，在进行最优一维搜索时，一般要先确定出问题的一个搜索区间，然后再在此区间中进行搜索求解.

确定搜索区间的一种简单方法叫进退法(advanced and retreat method)，其基本思想是从一点出发，按一定步长，试图确定出函数值呈现"高—低—高"的三个点，一个方向不成功，就退回来沿相反向搜索. 具体地说，就是给出初始点λ_0，初始步长$h_0>0$，若$\varphi(\lambda_0+h_0)<\varphi(\lambda_0)$，则下一步从新点$\lambda_0+h_0$出发，加大步长向前搜索，直到目标函数上升就停止. 若$\varphi(\lambda_0+h_0)>\varphi(\lambda_0)$，则下一步仍以$\lambda_0$为出发点，沿反方向搜索，直到目标函数上升就停止. 这样便得到一个搜索区间.

算法 6-1(进退法)

Step 1 选取初始数据. 给定初始点λ_0，初始步长$h_0>0$，加倍系数$\alpha>1$(一般取$\alpha=2$)，计算$\varphi_0=\varphi(\lambda_0)$，置$k=0,i=0$.

Step 2 试探. 令$\lambda_{k+1}=\lambda_k+h_k$，计算$\varphi_{k+1}=\varphi(\lambda_{k+1})$.

Step 3 比较目标函数值. 若$\varphi_{k+1}<\varphi_k$，转 Step 4，否则，转 Step 5.

Step 4 加步探索. 令

$$h_{k+1}=\alpha h_k,\lambda=\lambda_k,\lambda_k:=\lambda_{k+1},\varphi_k:=\varphi_{k+1},k:=k+1,$$

转 Step 2.

Step 5 反向探索. 若$k=0$且$i=0$，转换搜索方向，令$h_k:=-h_k,\lambda=\lambda_{k+1}$,

$i := i+1$，转 Step 2；否则，停止迭代. 令

$$a = \min\{\lambda, \lambda_{k+1}\}, b = \max\{\lambda, \lambda_{k+1}\}$$

输出搜索区间 $[a,b]$.

进退法的计算框图如图 6.1.2 所示.

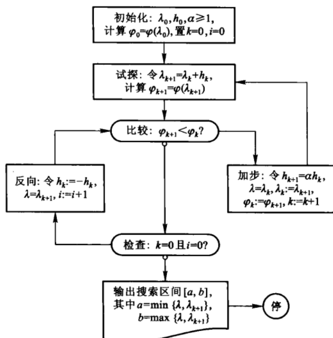


图 6.1.2　确定搜索区间的进退法计算框图

例 6.1.1　用进退法确定一维极小化问题

$$\min_{\lambda \geqslant 0} \varphi(\lambda) = \lambda^3 - 2\lambda + 1$$

的搜索区间，要求选取 $\lambda_0 = 0, h_0 = 1, \alpha = 2$.

解　Step 1　初始化：$\lambda_0 = 0, h_0 = 1, \varphi_0 = 1, k = 0$.

第一次搜索：

Step 2　$\lambda_1 = 1, \varphi_1 = 0$.

Step 3　$\varphi_1 < \varphi_0$，进行 Step 4.

Step 4　加步搜索. $h_1 = 2h_0, \lambda = 0, \lambda_0 = 1, \varphi_0 = 0, k = 1$，进入第二次搜索.

第二次搜索：

Step 2　$\lambda_2 = 3, \varphi_2 = 22$.

Step 3　$\varphi_2 > \varphi_1$，转 Step 5.

Step 5　$k \neq 0, a = \min\{\lambda, \lambda_2\} = 0, b = \max\{\lambda, \lambda_2\} = 3$，停止迭代，搜索区间

为$[0,3]$. □

图6.1.3给出了目标函数$\varphi(\lambda)$的图形.

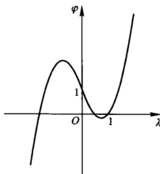


图6.1.3 函数$\varphi(\lambda)=\lambda^3-2\lambda+1$的图形

6.1.3 单谷函数及其性质

由于以后要介绍的一些一维搜索方法,主要针对单谷区间(或叫单峰区间,unimodal interval)和单谷函数(或叫单峰函数,unimodal function),这里介绍它们的概念及其简单性质.

定义6.1.2 设$\varphi:\mathbb{R}\to\mathbb{R}$,$[a,b]\subset\mathbb{R}$,若存在$\bar{\lambda}\in[a,b]$,使得$\varphi(\lambda)$在$[a,\bar{\lambda}]$上严格单调减小,在$[\bar{\lambda},b]$上严格单调增加,则称$[a,b]$是函数$\varphi(\lambda)$的单谷区间,$\varphi(\lambda)$是$[a,b]$上的单谷函数或单峰函数.

单谷函数也可按如下方式定义:

定义6.1.2a 如果存在$\bar{\lambda}\in[a,b]$,使得对于任何$\lambda_1,\lambda_2\in[a,b]$且$\lambda_1<\lambda_2$,当$\lambda_2<\bar{\lambda}$,有$\varphi(\lambda_1)>\varphi(\lambda_2)$;当$\lambda_1>\bar{\lambda}$,有$\varphi(\lambda_1)<\varphi(\lambda_2)$,则称$\varphi(\lambda)$是$[a,b]$上的单谷函数.

在单谷函数定义中,$\bar{\lambda}$是$\varphi(\lambda)$在$[a,b]$上的惟一极小点.图6.1.4给出了单谷函数和非单谷函数的图例.

定理6.1.2 设$\varphi:\mathbb{R}\to\mathbb{R}$,$[a,b]$为$\varphi(\lambda)$的单谷区间,$\lambda_1,\lambda_2\in[a,b]$,且$\lambda_1<\lambda_2$,那么

(1) 若$\varphi(\lambda_1)\leqslant\varphi(\lambda_2)$,则$[a,\lambda_2]$是$\varphi(\lambda)$的单谷区间;

(2) 若$\varphi(\lambda_1)\geqslant\varphi(\lambda_2)$,则$[\lambda_1,b]$是$\varphi(\lambda)$的单谷区间.

证明 设单谷函数$\varphi(\lambda)$在单谷区间$[a,b]$上的惟一极小点为$\bar{\lambda}$.要证(1),只要证$\bar{\lambda}\in[a,\lambda_2]$.若不然,$\lambda_1<\lambda_2<\bar{\lambda}$,则由定义6.1.2a知,$\varphi(\lambda_1)>\varphi(\lambda_2)$,此与(1)中的条件$\varphi(\lambda_1)\leqslant\varphi(\lambda_2)$矛盾.同理可证(2). □

根据上述定理,对于单谷函数,只需比较两个试探点的函数值,就可将包含

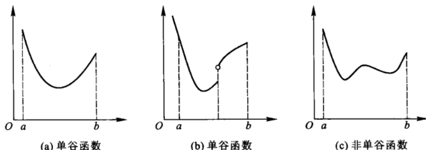


图 6.1.4　单谷函数与非单谷函数

极小点的区间缩短.

6.2　0.618 法和 Fibonacci 法

考虑一维无约束极小化问题

$$\min\ \varphi(t),t\in\mathbb{R}.$$

假定 $\varphi(t)$ 的一个搜索区间 $[a_1,b_1]$ 已确定,并设 $\varphi(t)$ 在 $[a_1,b_1]$ 上为单谷函数,$\bar{t}$ 为 $\varphi(t)$ 在 $[a_1,b_1]$ 上的极小点. 根据单谷函数的性质,只需增加两个试探点,就可把搜索区间缩短.

任取 $\lambda_1,\mu_1\in[a_1,b_1]$ 且 $\lambda_1<\mu_1$,当 $\varphi(\lambda_1)<\varphi(\mu_1)$ 时,取 $a_2=a_1,b_2=\mu_1$;当 $\varphi(\lambda_1)\geqslant\varphi(\mu_1)$ 时,取 $a_2=\lambda_1,b_2=b_1$. 从而由定理 6.1.2,总有 $\bar{t}\in[a_2,b_2]$. 继续上述过程,得一搜索区间序列

$$[a_1,b_1]\supseteq[a_2,b_2]\supseteq\cdots\supseteq[a_n,b_n],$$

且 $\bar{t}\in[a_n,b_n]$,当 $b_n-a_n<\varepsilon$ 时,$[a_n,b_n]$ 中任一点都可作为 $\bar{t}$ 的近似值.

那么,应该如何选择试探点呢?

在每次收缩区间时,新的搜索区间内已保留一个旧的试探点,因此只需再选一个新的试探点. 称

$$\tau_k=\frac{b_{k+1}-a_{k+1}}{b_k-a_k},k=1,2,\cdots$$

$(0<\tau_k<1)$ 为区间缩短率(interval constraction rate). 对 τ_k 有两种可能情况:$\tau_k=$ 常数或 $\tau_k\neq$ 常数. 相应地,我们分别有 0.618 法和 Fibonacci 法.

6.2.1　0.618 法(近似黄金分割法)

设在第 k 次迭代时,搜索区间为 $[a_k,b_k]$,为进一步缩短搜索区间,选取两个试探点 $\lambda_k,\mu_k\in[a_k,b_k]$,且 $\lambda_k<\mu_k$,则有

情形 1　当 $\varphi(\lambda_k)<\varphi(\mu_k)$ 时，$\bar{t}\in[a_k,\mu_k]=[a_{k+1},b_{k+1}]$.

情形 2　当 $\varphi(\lambda_k)\geqslant\varphi(\mu_k)$ 时，$\bar{t}\in[\lambda_k,b_k]=[a_{k+1},b_{k+1}]$.

现在要确定 λ_k,μ_k，使它们满足下述两个条件：

(1) λ_k 和 μ_k 在区间 $[a_k,b_k]$ 中的位置对称. 这样，不管删去哪一段，新的搜索区间 $[a_{k+1},b_{k+1}]$ 的长度不依赖于第 k 次迭代结果，即与 $\varphi(\lambda_k)<\varphi(\mu_k)$ 还是 $\varphi(\lambda_k)\geqslant\varphi(\mu_k)$ 无关.

于是有

$$b_{k+1}-a_{k+1}=\tau_k(b_k-a_k), \tag{6.2.1}$$

$$b_k-\lambda_k=\mu_k-a_k. \tag{6.2.2}$$

对于情形 1 和情形 2，式(6.2.1)分别对应于如下两式：

$$\mu_k-a_k=\tau_k(b_k-a_k), \tag{6.2.3}$$

$$b_k-\lambda_k=\tau_k(b_k-a_k). \tag{6.2.4}$$

由式(6.2.4)和(6.2.3)分别得到

$$\lambda_k=a_k+(1-\tau_k)(b_k-a_k), \tag{6.2.5}$$

$$\mu_k=a_k+\tau_k(b_k-a_k). \tag{6.2.6}$$

(2) 为减少计算量，在第 $k+1$ 次迭代中，保留一个旧的试探点，只增加一个新的试探点，如图 6.2.1 所示.

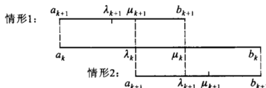


图 6.2.1　搜索区间与试探点的位置

对情形 1，由(6.2.6)和(6.2.1)两式，有

$$\begin{aligned}\lambda_k=\mu_{k+1}&=a_{k+1}+\tau_{k+1}(b_{k+1}-a_{k+1})\\&=a_k+\tau_{k+1}\tau_k(b_k-a_k).\end{aligned} \tag{6.2.7}$$

对情形 2，由(6.2.5)和(6.2.1)两式，有

$$\begin{aligned}\mu_k=\lambda_{k+1}&=a_{k+1}+(1-\tau_{k+1})(b_{k+1}-a_{k+1})\\&=a_k+(1-\tau_k)(b_k-a_k)+(1-\tau_{k+1})\tau_k(b_k-a_k)\\&=a_k+(1-\tau_{k+1}\tau_k)(b_k-a_k).\end{aligned} \tag{6.2.8}$$

比较(6.2.5)和(6.2.7)式以及(6.2.6)和(6.2.8)式得

$$\tau_{k+1}\tau_k=1-\tau_k. \tag{6.2.9}$$

这就是当试探点满足条件(1)和(2)时，τ_k 应满足的条件.

在要求每次迭代区间长度的缩短率相同，即 $\tau_k=\tau$（常数）的条件下，由(6.2.9)式得到

$$\tau^2+\tau-1=0, \tag{6.2.10}$$

解得

$$\tau=\frac{-1+\sqrt{5}}{2}\approx 0.618.$$

于是，由(6.2.5)和(6.2.6)两式得到了0.618法中试探点的迭代计算公式：

$$\begin{cases}\lambda_k=a_k+(1-\tau)(b_k-a_k),\\ \mu_k=a_k+\tau(b_k-a_k).\end{cases}$$

由于每次函数计算后搜索区间的缩短率为 τ，初始区间为 $[a_1,b_1]$，因此最终区间的长度为 $\tau^{n-1}(b_1-a_1)$，从而可知0.618法的收敛速度是线性的.

0.618法也叫近似黄金分割法(approximate golden section method)，这是因为缩短率 τ 恰为黄金分割数(golden section number)，它满足比例式

$$\frac{\tau}{1}=\frac{1-\tau}{\tau},$$

即满足方程(6.2.10). 其几何意义是：黄金分割数 τ 对应的点在单位长区间[0,1]中的位置相当于其对称点 $1-\tau$ 在区间 $[0,\tau]$ 中的位置（如图6.2.2所示）.

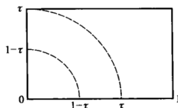


图6.2.2 黄金分割数的几何意义

算法6-2(0.618法)

Step 1 选取初始数据. 确定初始搜索区间 $[a_1,b_1]$ 和允许误差 $\varepsilon>0$，$\tau=0.618$.

Step 2 计算最初两个试探点：

$$\lambda_1=a_1+(1-\tau)(b_1-a_1),\ \mu_1=a_1+\tau(b_1-a_1),$$

求出 $\varphi(\lambda_1)$，$\varphi(\mu_1)$，并置 $k=1$.

Step 3 检查 $|\lambda_k-\mu_k|<\varepsilon$? 是，停止计算，输出 $\bar{t}=\dfrac{\lambda_k+\mu_k}{2}$；否则，转 Step 4.

Step 4 比较函数值. 若 $\varphi(\lambda_k)<\varphi(\mu_k)$，转 Step 5；若 $\varphi(\lambda_k)\geqslant\varphi(\mu_k)$，转

Step 6.

Step 5　向左搜索. 令

$$a_{k+1} := a_k, b_{k+1} = \mu_k, \mu_{k+1} = \lambda_k, \varphi(\mu_{k+1}) = \varphi(\lambda_k),$$

并计算

$$\lambda_{k+1} = a_{k+1} + (1-\tau)(b_{k+1} - a_{k+1}),$$

和 $\varphi(\lambda_{k+1})$，转 Step 7.

Step 6　向右搜索. 令

$$a_{k+1} = \lambda_k, b_{k+1} := b_k, \lambda_{k+1} = \mu_k, \varphi(\lambda_{k+1}) = \varphi(\mu_k),$$

并计算

$$\mu_{k+1} = a_{k+1} + \tau(b_{k+1} - a_{k+1})$$

和 $\varphi(\mu_{k+1})$，转 Step 7.

Step 7　置 $k := k+1$，转 Step 3.

在图 6.2.3 中，我们省去了算法 6－2 中的下标 k，给出了更接近计算机编程实现的算法框图.

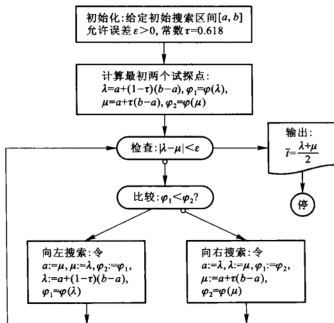


图 6.2.3　0.618 法计算框图

6.2.2 Fibonacci 法

另一种与 0.618 法相类似的分割法叫 Fibonacci 法. 它与 0.618 法的主要区别之一在于:搜索区间长度的缩短率不是采用黄金分割数,而是由所谓的 Fibonacci 数确定. Fibonacci 数是指满足下述条件的数列 $\{F_k\}$:

$$\begin{cases} F_0 = F_1 = 1, \\ F_{k+1} = F_k + F_{k-1}, k = 1,2,\cdots. \end{cases} \tag{6.2.11}$$

用数学归纳法可以证明,Fibonacci 数可由下式表示:

$$F_n = \frac{1}{\sqrt{5}}\left[\left(\frac{1+\sqrt{5}}{2}\right)^{n+1} - \left(\frac{1-\sqrt{5}}{2}\right)^{n+1}\right], n = 0,1,2,\cdots. \tag{6.2.12}$$

现在来看,按怎样的规则选取试探点可以使给定的搜索区间长度尽快地缩短? 换句话说,当事先给定计算试探点的函数值的总次数时,求能把多长的搜索区间缩短为长度为 1 的区间.

设计算试探点的函数值的总次数为 n,最终区间长度为 1,最初区间 $[a,b] = [a_0,b_0]$ 的长度为 L_n. 最优的选点方式应保证 L_n 最大.

设 L_i 的上确界为 $U_i(i=0,1,\cdots,n)$,由于至少要计算两次函数值才能缩短区间,故

$$U_0 = U_1 = 1, \tag{6.2.13}$$

且 $[a_1,b_1] = [a,b]$. 设最初的两个试探点为 λ_1,μ_1,且 $\lambda_1 < \mu_1$,那么余下还可计算函数值 $n-2$ 次,极小点可能位于 $[a_1,\lambda_1]$ 或 $[\lambda_1,b_1]$ 中. 当极小点位于 $[a_1,\lambda_1]$ 中时,必须再计算 $n-2$ 次才能将区间缩短为 1,因此

$$\lambda_1 - a_1 \leqslant U_{n-2}. \tag{6.2.14}$$

当极小点位于 $[\lambda_1,b_1]$ 中时,除了可再计算函数值 $n-2$ 次外,还可利用已算出的函数值 $\varphi(\mu_1)$,因此总共可利用 $n-1$ 次函数值,所以

$$b_1 - \lambda_1 \leqslant U_{n-1}. \tag{6.2.15}$$

由(6.2.14)及(6.2.15)两式可得

$$L_n = b - a = b_1 - a_1 \leqslant U_{n-1} + U_{n-2}.$$

从而

$$U_n \leqslant U_{n-1} + U_{n-2}, \tag{6.2.16}$$

并且由(6.2.11)和(6.2.16)两式可得 $U_n \leqslant F_n$. 因此,如果能保证计算 n 次函数值后,可把长度是 F_n 的最初区间缩短为长度是 1 的区间,或即把搜索区间长度缩短为最初区间长度的 $\frac{1}{F_n}$,那么就可以认为这种选点方式最优. 按这种最优选点方式,可以取搜索区间长度的缩短率

$$\tau_k = \frac{b_{k+1} - a_{k+1}}{b_k - a_k} = \frac{F_{n-k}}{F_{n-k+1}}.$$

容易验证(6.2.9)式成立,即 $\tau_{k+1}\tau_k = 1 - \tau_k$. 根据(6.2.5)及(6.2.6)式,得到 Fibonacci 法试探点迭代计算公式

$$\begin{cases}\lambda_k = a_k + \dfrac{F_{n-k-1}}{F_{n-k+1}}(b_k - a_k), k = 1,2,\cdots,n-1,\\ \mu_k = a_k + \dfrac{F_{n-k}}{F_{n-k+1}}(b_k - a_k), k = 1,2,\cdots,n-1.\end{cases} \tag{6.2.17}$$

如果给定最终区间的长度不超过 $\varepsilon > 0$,即 $b_n - a_n \leqslant \varepsilon$,根据

$$b_n - a_n = \frac{F_1}{F_2}(b_{n-1} - a_{n-1}) = \frac{F_1}{F_2}\cdot\frac{F_2}{F_3}\cdot\cdots\cdot\frac{F_{n-1}}{F_n}(b_1 - a_1)$$

$$= \frac{F_1}{F_n}(b_1 - a_1) = \frac{1}{F_n}(b - a)$$

有

$$F_n \geqslant \frac{b-a}{\varepsilon}, \tag{6.2.18}$$

那么,由(6.2.18)式就可以确定出 n.

运用 Fibonacci 法,应注意下列问题:

由于第 1 次迭代计算两个试探点,以后每次计算一个,这样经过 $n-1$ 次迭代就计算完 n 个试探点.但是,根据公式(6.2.17)知

$$\lambda_{n-1} = a_{n-1} + \frac{F_0}{F_2}(b_{n-1} - a_{n-1}) = \frac{a_{n-1} + b_{n-1}}{2},$$

$$\mu_{n-1} = a_{n-1} + \frac{F_1}{F_2}(b_{n-1} - a_{n-1}) = \frac{a_{n-1} + b_{n-1}}{2}.$$

这时 $\lambda_{n-1} = \mu_{n-1}$,这说明在第 $n-1$ 次迭代中并没有选择新的试探点.为了在第 $n-1$ 次迭代中能够缩短搜索区间,可令

$$\lambda_n = \lambda_{n-1}, \mu_n = \lambda_{n-1} + \delta,$$

其中辨别常数 $\delta > 0$.

算法 6-3(Fibonacci 法)

Step 1　选取初始数据.给定初始搜索区间 $[a_1, b_1]$ 和允许误差 $\varepsilon > 0$,辨别系数 $\delta > 0$,求搜索次数 n,使

$$F_n \geqslant \frac{b_1 - a_1}{\varepsilon}.$$

Step 2　计算最初两个试探点:

$$\lambda_1 = a_1 + \frac{F_{n-2}}{F_n}(b_1 - a_1), \mu_1 = a_1 + \frac{F_{n-1}}{F_n}(b_1 - a_1).$$

求函数值 $\varphi(\lambda_1)$ 和 $\varphi(\mu_1)$,置 $k=1$.

Step 3　检查 $\varphi(\lambda_k) < \varphi(\mu_k)$? 是,转 Step 4;否,转 Step 5.

Step 4　向左搜索. 令

$$a_{k+1} := a_k, b_{k+1} = \mu_k, \mu_{k+1} = \lambda_k, \varphi(\mu_{k+1}) = \varphi(\lambda_k),$$

并计算

$$\lambda_{k+1} = a_{k+1} + \frac{F_{n-k-2}}{F_{n-k}}(b_{k+1} - a_{k+1})$$

和 $\varphi(\lambda_{k+1})$,转 Step 6.

Step 5　向右搜索. 令

$$a_{k+1} = \lambda_k, b_{k+1} := b_k, \lambda_{k+1} = \lambda_k, \varphi(\lambda_{k+1}) = \varphi(\mu_k),$$

并计算

$$\mu_{k+1} = a_{k+1} + \frac{F_{n-k-1}}{F_{n-k}}(b_{k+1} - a_{k+1})$$

和 $\varphi(\mu_{k+1})$,转 Step 6.

Step 6　置 $k := k+1$,若 $k < n-1$,转 Step 3;若 $k = n-1$,转 Step 7.

Step 7　令 $\lambda_n = \lambda_{n-1}, \mu_n = \lambda_{n-1} + \delta$,计算 $\varphi(\lambda_n)$ 和 $\varphi(\mu_n)$. 若 $\varphi(\lambda_n) < \varphi(\mu_n)$,则令 $a_n := a_{n-1}, b_n = \mu_n$;否则,令 $a_n = \lambda_n, b_n := b_{n-1}$,停止计算,极小点含于区间$[a_n, b_n]$.

前面已经证明了 Fibonacci 法是分割方法求一维极小化问题的最优策略. 根据(6.2.12)式知

$$\lim_{k\to\infty} \frac{F_{k-1}}{F_k} = \frac{\sqrt{5}-1}{2} = \tau.$$

这表明,当 $k\to\infty$ 时,Fibonacci 法与 0.618 法的区间缩短率相同,因而 Fibonacci 法也以收敛比 τ 线性收敛;并且当 $k\to\infty$ 时 0.618 法也是最优策略. 比较而言,0.618 法简单易行,因而应用更为广泛.

6.3　函数逼近法

函数逼近法又称作插值法,它的基本思想是根据目标函数 $\varphi(t)$ 在某些点的信息,构造一个与它近似的多项式函数 $\psi(t)$,在一定条件下,可以期望 $\psi(t)$ 的极小点接近 $\varphi(t)$ 的极小点.

6.3.1　Newton 法(或 Newton 切线法)

Newton 法的基本思想是,在极小点附近用二阶 Taylor 多项式去近似目标函数 $\varphi(t)$,进而求出极小点的估计值.

令

$$\psi(t) = \varphi(t_k) + \varphi'(t_k)(t - t_k) + \frac{1}{2}\varphi''(t_k)(t - t_k)^2,$$

由 $\psi'(t) = \varphi'(t_k) + \varphi''(t_k)(t - t_k) = 0$,解得 $\psi(t)$ 的平稳点,记作 t_{k+1},则

$$t_{k+1} = t_k - \frac{\varphi'(t_k)}{\varphi''(t_k)}. \tag{6.3.1}$$

在点 t_k 附近,$\varphi(t) \approx \psi(t)$. 因此可用 $\psi(t)$ 的极小点作为目标函数 $\varphi(t)$ 的极小点的估计. (6.3.1)式亦可写作

$$t_{k+1} = t_k - [\varphi''(t_k)]^{-1}\varphi'(t_k). \tag{6.3.2}$$

(6.3.1)式或(6.3.2)式就是 Newton 法的迭代公式.

在微积分中,求函数 $\varphi(t)$ 的极小点可以用 $\varphi(t)$ 的平稳点,即方程

$$\varphi'(t) = 0 \tag{6.3.3}$$

的根来近似. 这个非线性方程的求解可以用 Newton 切线法来求解,其想法是用迭代点 t_k 处的切线

$$y - \varphi'(t_k) = \varphi''(t_k)(t - t_k)$$

与横轴交点的横坐标 t_{k+1} 来作为方程(6.3.3)的根的新的近似(如图 6.3.1). 令 $y = 0$ 得到(6.3.1)式. 因此 Newton 法又叫 Newton 切线法(Newton tangent method).

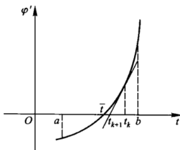


图 6.3.1 Newton 切线法

算法 6-4(Newton 法)

Step 1 给定初始点 t_0,允许误差 $\varepsilon > 0$,置 $k = 0$.

Step 2 检查 $|\varphi'(t_k)| < \varepsilon$? 是,输出 t_k,停止计算;否,转 Step 3.

Step 3 计算点

$$t_{k+1} = t_k - \frac{\varphi'(t_k)}{\varphi''(t_k)},$$

置 $k := k + 1$,转 Step 2.

下述定理表明 Newton 法具有局部二阶收敛速度.

定理 6.3.1 设 $\varphi(t)$ 存在三阶连续导数，$\bar{t}$ 满足

$$\varphi'(\bar{t}) = 0, \varphi''(\bar{t}) \neq 0,$$

则当初始点 t_0 充分接近 $\bar{t}$ 时，由 Newton 法迭代产生的点列 $\{t_k\}$ 至少二阶收敛于 $\bar{t}$.

证明 设 $t_k \neq \bar{t}, k = 1,2,\cdots$，注意到 $\varphi'(\bar{t}) = 0$，由 Newton 法迭代公式有

$$\begin{aligned}
|t_{k+1} - \bar{t}| &= \left| t_k - \frac{\varphi'(t_k)}{\varphi''(t_k)} - \bar{t} \right| \\
&= \frac{1}{|\varphi''(t_k)|} |t_k\varphi''(t_k) - \varphi'(t_k) - \bar{t}\varphi''(t_k)| \\
&= \frac{1}{|\varphi''(t_k)|} |\varphi'(\bar{t}) - [\varphi'(t_k) + \varphi''(t_k)(\bar{t} - t_k)]| \\
&= \frac{1}{|\varphi''(t_k)|} \cdot \frac{1}{2}(\bar{t} - t_k)^2 |\varphi'''(\xi_k)|, \qquad (6.3.4)
\end{aligned}$$

其中 ξ_k 介于 $\bar{t}$ 和 t_k 之间.

因 $\varphi''(t), \varphi'''(t)$ 连续，$\varphi''(\bar{t}) \neq 0$，因此，当 t_k 充分接近于 $\bar{t}$ 时，必有 $M_1 > 0$，$M_2 > 0$，使得在包含 $\bar{t}$ 和 t_k 的闭区间的每点 t 上，都有

$$|\varphi''(t)| \geq M_1, |\varphi'''(t)| \leq M_2.$$

从而由(6.3.4)式知

$$|t_{k+1} - \bar{t}| \leq \frac{M_2}{2M_1}(t_k - \bar{t})^2. \qquad (6.3.5)$$

取初始点 t_0 充分接近于 $\bar{t}$，使得

$$\frac{M_2}{M_1}|t_0 - \bar{t}| < 1.$$

于是有

$$\{t_k\} \subset X = \{t \mid |t - \bar{t}| \leq |t_0 - \bar{t}|\},$$

且

$$|t_{k+1} - \bar{t}| < \frac{1}{2}|t_k - \bar{t}|.$$

故易知 $\lim\limits_{k\to\infty}|t_k - \bar{t}| = 0$，即 $\lim\limits_{k\to\infty} t_k = \bar{t}$，且由(6.3.5)知 $\{t_k\}$ 的收敛阶为 2. □

Newton 法具有二阶局部收敛性，收敛速度快，当初始点靠近最优点时，通常经几次迭代就可得到满足精度要求的点. 但因为 Newton 法不具有全局收敛性，所以当初始点远离极小点时，迭代序列可能不收敛于极小点. 下面用一个数值例子加以说明.

例 6.3.1 用 Newton 法求解目标函数为

$$\varphi(t) = \begin{cases} 4t^3 - 3t^4, t \geq 0, \\ 4t^3 + 3t^4, t < 0 \end{cases}$$

的一维无约束极小化问题.

我们看到,当选取初始点 $t=0.4$ 时,如表 6.3.1 所示,6 次迭代后得到点 0.002 807,很快接近于稳定点 $t=0$. 当选取初始点 $t=0.6$ 时,如表 6.3.2 所示,迭代过程在点 0.6 和 -0.6 之间摆动. □

表 6.3.1 选取初始点 $t=0.4$ 时的迭代过程

迭代 k	t_k	$\varphi'(t_k)$	$\varphi''(t_k)$	t_{k+1}
1	0.400 000	1.152 000	3.840 000	0.100 000
2	0.100 000	0.108 000	2.40 000	0.047 059
3	0.047 059	0.025 324	1.049 692	0.022 934
4	0.022 934	0.006 167	0.531 481	0.011 331
5	0.011 331	0.001 523	0.267 322	0.005 634
6	0.005 634	0.000 379	0.134 073	0.002 807

表 6.3.2 选取初始点 $t=0.6$ 时的迭代过程

迭代 k	t_k	$\varphi'(t_k)$	$\varphi''(t_k)$	t_{k+1}
1	0.600	1.728	1.440	-0.600
2	-0.600	1.728	-1.440	0.600
3	0.600	1.728	1.440	-0.600
4	-0.600	1.728	-1.440	0.600

图 6.3.2 给出例 6.3.1 中函数 $\varphi(t)$ 的图形.

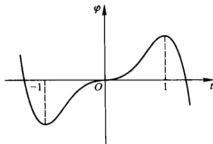


图 6.3.2 例 6.3.1 中函数 $\varphi(t)$ 的图形

6.3.2 割线法

割线法(secant method)的基本思想是用区间$[t_{k-1},t_k]$(或$[t_k,t_{k-1}]$)上的割线近似代替目标函数的导函数的曲线,并用该割线与横轴交点的横坐标作为方程(6.3.3)式的根的近似.

如图 6.3.3 所示. 区间$[t_{k-1},t_k]$(或$[t_k, t_{k-1}]$)上的割线方程为

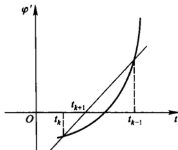


图 6.3.3 割线法

$$y-\varphi'(t_k)=\frac{\varphi'(t_k)-\varphi'(t_{k-1})}{t_k-t_{k-1}}(t-t_k),$$

令 $y=0$,得到割线法的迭代公式

$$t_{k+1}=t_k-\frac{t_k-t_{k-1}}{\varphi'(t_k)-\varphi'(t_{k-1})}\varphi'(t_k). \tag{6.3.6}$$

仿照 Newton 法的算法步骤可以给出割线法的算法步骤.

割线法具有超线性局部收敛性,与 Newton 法相比,收敛速度较慢,但不需要计算二阶导数.

引理 6.3.2 设 $a<b<c$,函数 $\varphi(t)$在$[a,c]$上具有二阶导数,则至少存在一点 $\xi\in[a,c]$,使得

$$\frac{\varphi(a)}{(a-b)(a-c)}+\frac{\varphi(b)}{(b-a)(b-c)}+\frac{\varphi(c)}{(c-a)(c-b)}=\frac{1}{2}\varphi''(\xi).$$

证明留作练习. □

定理 6.3.3 设 $\varphi(t)$存在连续的三阶导数,$\bar{t}$ 满足 $\varphi'(\bar{t})=0,\varphi''(\bar{t})\neq0$. 若 t_1 和 t_2 充分接近 $\bar{t}$,则由割线法产生的迭代点列$\{t_k\}$收敛于 $\bar{t}$,其收敛速度的阶为 $\tau=\dfrac{1+\sqrt{5}}{2}\approx1.618$.

证明 因 $\varphi''(t)$连续且 $\varphi''(\bar{t})\neq0$,则存在一个包含 $\bar{t}$ 的充分小的闭区间 $\Delta=\{t\mid |t-\bar{t}|\leqslant\delta\}$,使得对每一个 $t\in\Delta$,都有 $\varphi''(t)\neq0$. 任取 $t_1,t_2\in\Delta$,以 $t_k,t_{k-1}\in\Delta$ 为节点构造插值多项式

$$\psi(t)=\varphi'(t_k)+\frac{\varphi'(t_k)-\varphi'(t_{k-1})}{t_k-t_{k-1}}(t-t_k), \tag{6.3.7}$$

其插值余项

$$\varphi'(t)-\psi(t)=\left[\frac{\varphi'(t_k)}{(t_k-t_{k-1})(t_k-t)}+\frac{\varphi'(t)}{(t-t_k)(t-t_{k-1})}+\frac{\varphi'(t_{k-1})}{(t_{k-1}-t_k)(t_{k-1}-t)}\right](t-t_{k-1})(t-t_k).$$

根据引理 6.3.2 知

$$\varphi'(t)-\psi(t)=\frac{1}{2}\varphi'''(\xi_1)(t-t_{k-1})(t-t_k).$$

其中 ξ_1 介于 t_k 与 t_{k-1} 之间，且 $\xi_1\in\Delta$.

记 $\eta_k=t_k-\bar{t}$，并注意到 $\varphi'(\bar{t})=0$，有

$$\psi(\bar{t})=-\frac{1}{2}\varphi'''(\xi_1)\eta_k\eta_{k-1}. \tag{6.3.8}$$

由(6.3.6)及(6.3.7)两式知

$$\psi(t_{k+1})=0,\psi'(t)=\frac{\varphi'(t_k)-\varphi'(t_{k-1})}{t_k-t_{k-1}}.$$

因此，有

$$\begin{aligned}\psi(\bar{t})&=\psi(\bar{t})-\psi(t_{k+1})=\psi'(\xi_2)(\bar{t}-t_{k+1})\\&=\frac{\varphi'(t_k)-\varphi'(t_{k-1})}{t_k-t_{k-1}}(\bar{t}-t_{k+1})\\&=-\varphi''(\xi_3)(t_{k+1}-\bar{t})=-\varphi''(\xi_3)\eta_{k+1}.\end{aligned} \tag{6.3.9}$$

其中 ξ_2 介于 $\bar{t}$ 与 t_{k+1} 之间，ξ_3 介于 t_k 与 t_{k-1} 之间，且 $\xi_3\in\Delta$.

比较(6.3.8)与(6.3.9)两式得到

$$\eta_{k+1}=\frac{\varphi'''(\xi_1)}{2\varphi''(\xi_3)}\eta_k\eta_{k-1}. \tag{6.3.10}$$

令 $M=\dfrac{\max\limits_{t\in\Delta}|\varphi'''(t)|}{2\min\limits_{t\in\Delta}|\varphi''(t)|}$，得

$$|\eta_{k+1}|\leqslant M|\eta_k|\cdot|\eta_{k-1}|. \tag{6.3.11}$$

取充分小的 Δ，使得 $M\delta<1$，则有

$$|\eta_{k+1}|\leqslant M\delta\delta<\delta.$$

这说明由 $t_k,t_{k-1}\in\Delta$ 必可推得 $t_{k+1}\in\Delta$，进而由 $t_1,t_2\in\Delta$ 推知 $\{t_k\}\in\Delta$，且由(6.3.11)式可知

$$|\eta_{k+1}|\leqslant M\delta|\eta_k|,$$

所以易知 $\{t_k\}$ 收敛于 $\bar{t}$.

下面来讨论点列 $\{t_k\}$ 的收敛速度.

注意到 φ'' 和 φ''' 连续，则当 $k\to\infty$ 时，

$$\frac{\varphi'''(\xi_1)}{2\varphi''(\xi_3)}\to\frac{1}{2}\frac{\varphi'''(\bar{t})}{\varphi''(\bar{t})}.$$

因此，当 k 充分大时，由(6.3.10)式得

$$|\eta_{k+1}|\approx\bar{M}|\eta_k|\cdot|\eta_{k-1}|, \tag{6.3.12}$$

其中 $\bar{M}=\dfrac{1}{2}\dfrac{|\varphi'''(\bar{t})|}{|\varphi''(\bar{t})|}$. 记 $\bar{M}|\eta_k|=e^{u_k}$，则由(6.3.12)式知

$$u_{k+1} = u_k + u_{k-1},$$

它是 Fibonacci 数列所满足的方程. 这个差分方程的特征方程 $\tau^2 - \tau - 1 = 0$ 对应的两个特征根为

$$\tau_1 = \frac{1+\sqrt{5}}{2}, \tau_2 = \frac{1-\sqrt{5}}{2}.$$

这样,有

$$u_k = A\tau_1^k + B\tau_2^k, k = 1,2,\cdots,$$

其中 A 和 B 是待定的常数. 于是,我们得到

$$|\eta_k| = \frac{1}{\bar{M}}e^{A\tau_1^k + B\tau_2^k} \approx \frac{1}{\bar{M}}e^{A\tau_1^k}(k \gg 1), |\eta_{k+1}| \approx \frac{1}{\bar{M}}e^{A\tau_1^{k+1}}.$$

从而

$$\frac{|\eta_{k+1}|}{|\eta_k|^{\tau_1}} \approx \frac{\frac{1}{\bar{M}}e^{A\tau_1^{k+1}}}{\frac{1}{\bar{M}^{\tau_1}}e^{A\tau_1^{k+1}}} = \bar{M}^{\tau_1 - 1}.$$

所以,割线法的收敛阶是 $\tau_1 = \frac{1+\sqrt{5}}{2} \approx 1.618$. □

6.3.3 二次插值法

二次插值法(quadratic interpolation)又叫抛物线法(parabolic method),其基本思想是在极小点附近,用二次三项式 $\psi(t)$ 逼近目标函数 $\varphi(t)$. 设 $[a_k,b_k]$ 为第 k 次迭代的搜索区间,$t_k \in (a_k,b_k)$,今在 $[a_k,b_k]$ 上取二次插值多项式

$$\psi(t) = \alpha t^2 + \beta t + \gamma, \quad (6.3.13)$$

要求在插值点 a_k, b_k 和 t_k 处,$\psi(t)$ 与 $\varphi(t)$ 有相同的函数值(如图 6.3.4). 然后用 $\psi(t)$ 在 $[a_k,b_k]$ 上的极小点 t_{k+1} 作为 $\varphi(t)$ 在 $[a_k,b_k]$ 上的极小点的估计.

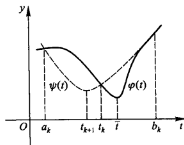


图 6.3.4 二次插值法

由三点插值条件,我们有

$$\begin{cases} \psi(a_k) = \alpha a_k^2 + \beta a_k + \gamma = \varphi(a_k), \\ \psi(t_k) = \alpha t_k^2 + \beta t_k + \gamma = \varphi(t_k), \\ \psi(b_k) = \alpha b_k^2 + \beta b_k + \gamma = \varphi(b_k). \end{cases}$$

由 Cramer 法则,解得

$$\alpha = \begin{vmatrix} \varphi(a_k) & a_k & 1 \\ \varphi(t_k) & t_k & 1 \\ \varphi(b_k) & b_k & 1 \end{vmatrix} \Bigg/ \begin{vmatrix} a_k^2 & a_k & 1 \\ t_k^2 & t_k & 1 \\ b_k^2 & b_k & 1 \end{vmatrix}$$

$$= -\frac{(b_k - t_k)\varphi(a_k) + (a_k - b_k)\varphi(t_k) + (t_k - a_k)\varphi(b_k)}{(b_k - t_k)(a_k - b_k)(t_k - a_k)}, \quad (6.3.14)$$

$$\beta = \begin{vmatrix} a_k^2 & \varphi(a_k) & 1 \\ t_k^2 & \varphi(t_k) & 1 \\ b_k^2 & \varphi(b_k) & 1 \end{vmatrix} \Bigg/ \begin{vmatrix} a_k^2 & a_k & 1 \\ t_k^2 & t_k & 1 \\ b_k^2 & b_k & 1 \end{vmatrix}$$

$$= \frac{(b_k^2 - t_k^2)\varphi(a_k) + (a_k^2 - b_k^2)\varphi(t_k) + (t_k^2 - a_k^2)\varphi(b_k)}{(b_k - t_k)(a_k - b_k)(t_k - a_k)}. \quad (6.3.15)$$

又由(6.3.13)式,令

$$\psi'(t) = 2\alpha t + \beta = 0,$$

得 $\psi(t)$ 的极小点

$$t_{k+1} = -\frac{\beta}{2\alpha}. \quad (6.3.16)$$

将(6.3.14)和(6.3.15)两式代入(6.3.16)式得迭代公式

$$t_{k+1} = \frac{1}{2}\frac{(b_k^2 - t_k^2)\varphi(a_k) + (a_k^2 - b_k^2)\varphi(t_k) + (t_k^2 - a_k^2)\varphi(b_k)}{(b_k - t_k)\varphi(a_k) + (a_k - b_k)\varphi(t_k) + (t_k - a_k)\varphi(b_k)}. \quad (6.3.17)$$

这个公式也可直接利用拉格朗日插值公式

$$L(t) = \frac{(t - t_k)(t - b_k)}{(a_k - t_k)(a_k - b_k)}\varphi(a_k) + \frac{(t - a_k)(t - b_k)}{(t_k - a_k)(t_k - b_k)}\varphi(t_k) + \frac{(t - a_k)(t - t_k)}{(b_k - a_k)(b_k - t_k)}\varphi(b_k),$$

并令 $L'(t) = 0$ 得到.

从初始的搜索区间$[a,b]$开始,反复利用公式(6.3.17),不断用二次插值多项式的极小点来缩短搜索区间,当搜索区间缩短到足够小时,便得到一维无约束极小化问题的近似最优解.

算法 6-5(二次插值法)

Step 1　选取初始数据. 确定初始搜索区间$[a,b]$,选取初始插值内点 $t_0 \in (a,b)$,给出允许误差 $\varepsilon > 0$,令

$$a_0 = a, b_0 = b, k = 0.$$

Step 2　求插值多项式的极小点

$$t_{k+1} = \frac{1}{2}\frac{(b_k^2 - t_k^2)\varphi(a_k) + (a_k^2 - b_k^2)\varphi(t_k) + (t_k^2 - a_k^2)\varphi(b_k)}{(b_k - t_k)\varphi(a_k) + (a_k - b_k)\varphi(t_k) + (t_k - a_k)\varphi(b_k)}.$$

Step 3　比较目标函数值. 若 $\varphi(t_{k+1}) \leqslant \varphi(t_k)$,则当 $|t_{k+1} - t_k| \leqslant \varepsilon$ 时,停止计算,输出 t_{k+1},否则,转 Step 4;若 $\varphi(t_{k+1}) > \varphi(t_k)$,则当 $|t_{k+1} - t_k| \leqslant \varepsilon$ 时,停止计算,输出 t_k,否则,转 Step 5;

Step 4　迭代点比插值点好时,确定新插值点. 检查迭代点 t_{k+1} 和插值点 t_k 的位置,若 $t_{k+1} \leqslant t_k$,令

$$a_{k+1} := a_k, b_{k+1} = t_k, k := k+1,$$

转 Step 2;否则,令

$$a_{k+1} = t_k, b_{k+1} := b_k, k := k+1,$$

转 Step 2.

Step 5　迭代点比插值点坏时,确定新插值点. 检查迭代点 t_{k+1} 和插值点 t_k 的位置,若 $t_{k+1} \leqslant t_k$,令

$$a_{k+1} = t_{k+1}, t_{k+1} := t_k, b_{k+1} := b_k, k := k+1,$$

转 Step 2;否则,令

$$a_{k+1} := a_k, b_{k+1} = t_{k+1}, t_{k+1} := t_k, k := k+1,$$

转 Step 2.

从上述算法步骤中可以看出,由于保证了

$$a_k < t_k < b_k, \varphi(a_k) > \varphi(t_k) < \varphi(b_k),$$

因此保证了插值多项式 $\psi(t)$ 的极小点 t_{k+1} 在区间 $[a_k, b_k]$ 中.

文献[50]证明了二次插值法的迭代公式(6.3.17)式产生的迭代点列 $\{t_k\}$ 收敛,且收敛阶为 1.32. 关于三次插值法和其他插值法也参见文献[50].

6.4　非精确一维搜索

前面我们介绍了一维精确搜索方法,通过求解

$$\min_{\lambda \geqslant 0} \varphi(\lambda) = f(\boldsymbol{x}_k + \lambda \boldsymbol{d}_k)$$

来获得从 $\boldsymbol{x}_k$ 出发沿 $\boldsymbol{d}_k$ 方向的最优步长 λ_k,并得到后继迭代点 $\boldsymbol{x}_{k+1} = \boldsymbol{x}_k + \lambda_k \boldsymbol{d}_k$. 这种精确搜索往往计算量很大,特别是当迭代点远离最优解时,效率很低. 而且,很多最优化算法的收敛速度并不依赖于精确一维搜索过程. 所以,我们应着眼于保证目标函数在每次迭代有满意的下降量的方法,这就是非精确一维搜索(inexact one dimension search)方法或可接受一维搜索(acceptable one dimension search)方法,它可以大大节省计算量.

6.4.1 Armijo - Goldstein 非精确一维搜索方法

设

$$J = \{\lambda > 0 \mid f(\boldsymbol{x}_k + \lambda \boldsymbol{d}_k) < f(\boldsymbol{x}_k)\}$$

是一个区间. 在图 6.4.1 中,区间 $J=(0,a)$. 为保证目标函数有一个满意的下降量,必须避免步长 λ 太靠近区间 J 的端点.

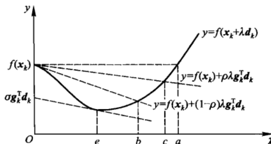


图 6.4.1 非精确一维搜索的含义

记

$$s_k = f(\boldsymbol{x}_k) - f(\boldsymbol{x}_k + \lambda_k \boldsymbol{d}_k),$$

它表示 f 在 $\boldsymbol{x}_k$ 处沿方向 $\boldsymbol{d}_k$ 的一个下降量,当我们控制 s_k 到一个满意水平时,λ_k 就是可接受步长.

因 $\boldsymbol{d}_k$ 是一个下降方向,记 $\boldsymbol{g}_k = \nabla f(\boldsymbol{x}_k)$,则

$$\varphi'(0) = \boldsymbol{g}_k^{\mathrm{T}} \boldsymbol{d}_k < 0.$$

于是,我们能给出 s_k 的一个合理的下界:

$$s_k \geqslant -\rho \lambda_k \boldsymbol{g}_k^{\mathrm{T}} \boldsymbol{d}_k, \tag{6.4.1}$$

其中 $\rho \in (0,1)$ 为常数. (6.4.1)式等价于

$$f(\boldsymbol{x}_k + \lambda_k \boldsymbol{d}_k) \leqslant f(\boldsymbol{x}_k) + \rho \lambda_k \boldsymbol{g}_k^{\mathrm{T}} \boldsymbol{d}_k. \tag{6.4.2}$$

满足(6.4.2)式要求的 λ_k 构成了如图 6.4.1 的区间 $J_1=(0,c]$,这排斥了区间 J 右端附近的点. 但是,它没有排除 λ_k 取接近于 0 的值. 而当 λ_k 接近于 0 时,一维搜索只能取得很少的进展,只有很小的下降量. 为了避免 λ_k 过小的情况,我们给 s_k 一个上界限制如下:

$$s_k \leqslant -(1-\rho)\lambda_k \boldsymbol{g}_k^{\mathrm{T}} \boldsymbol{d}_k,$$

它等价于

$$f(\boldsymbol{x}_k + \lambda_k \boldsymbol{d}_k) \geqslant f(\boldsymbol{x}_k) + (1-\rho)\lambda_k \boldsymbol{g}_k^{\mathrm{T}} \boldsymbol{d}_k. \tag{6.4.3}$$

这个要求排斥了区间 J 的左端点附近的点. 满足(6.4.2)和(6.4.3)两式要求的 λ_k 构成了区间 $J_2=[b,c]$. 当然,此时要求 $0<\rho<\frac{1}{2}$. 我们把(6.4.2)和(6.4.3)

式称为 Armijo - Goldstein 非精确线性搜索准则(Armijo - GoldStein inexact line search criterion),简称 Armijo - Goldstein 准则,它是由 Armijo(1966)和 Goldstein(1965)分别提出的. 把满足(6.4.2)和(6.4.3)两式要求的区间 $J_2=[b,c]$ 称为可接受区间,J_2 中的点称为可接受步长.

式(6.4.2)和式(6.4.3)也可分别写成

$$\varphi(\lambda_k) \leqslant \varphi(0) + \rho\lambda_k\varphi'(0), \tag{6.4.2a}$$

$$\varphi(\lambda_k) \geqslant \varphi(0) + (1-\rho)\lambda_k\varphi'(0). \tag{6.4.3a}$$

在几何上,它表示可接受步长 λ_k 是使 $\varphi(\lambda_k)$ 位于两直线 $y=\varphi(0)+\rho\lambda_k\varphi'(0)$ 和 $y=\varphi(0)+(1-\rho)\lambda\varphi'(0)$ 之间的点.

算法 6-6(Armijo - Goldstein 非精确一维搜索算法)

Step 1 选取初始数据. 在搜索区间$[0,\infty)$(或$[0,\lambda_{\max}]$)中取定初始点 λ_0,计算 $\varphi(0)$ 和 $\varphi'(0)$,给出可接受系数 $\rho\in\left(0,\frac{1}{2}\right)$,增大试探点系数 $\alpha>1$,令 $a_0=0, b_0=\infty$(或 $\lambda_{\max}$),$k=0$.

Step 2 检查下降量要求. 计算 $\varphi(\lambda_k)$,若

$$\varphi(\lambda_k) \leqslant \varphi(0) + \rho\lambda_k\varphi'(0),$$

转 Step 3;否则,令

$$a_{k+1} := a_k, b_{k+1} = \lambda_k,$$

转 Step 4.

Step 3 检查避免探索点过小要求. 若

$$\varphi(\lambda_k) \geqslant \varphi(0) + (1-\rho)\lambda_k\varphi'(0),$$

停止迭代,输出 λ_k,否则,令

$$a_{k+1} = \lambda_k, b_{k+1} := b_k,$$

若 $b_{k+1}<\infty$,转 Step 4,否则,令

$$\lambda_{k+1} = \alpha\lambda_k, k := k+1,$$

转 Step 2.

Step 4 选取新的试探点. 取

$$\lambda_{k+1} = \frac{a_{k+1}+b_{k+1}}{2}$$

令 $k:=k+1$,转 Step 2.

图 6.4.2 给出了该算法的计算框图.

6.4.2 Wolfe - Powell 非精确一维搜索方法

如图 6.4.1 所示,Armijo - Goldsteim 准则有可能把最优步长排除在可接受区间外面,为此,Wolfe - Powell 给出了一个更简单的条件代替(6.4.3a)式:

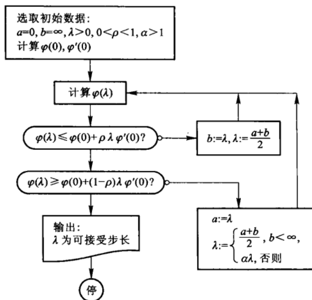


图 6.4.2 Armijo - Goldstein 非精确一维搜索算法框图

$$\boldsymbol{g}_{k+1}^{\mathrm{T}}\boldsymbol{d}_k \geqslant \sigma \boldsymbol{g}_k^{\mathrm{T}}\boldsymbol{d}_k, \sigma \in (\rho, 1), \tag{6.4.4}$$

亦即

$$\begin{aligned}\varphi'(\lambda_k) &= \nabla f(\boldsymbol{x}_k + \lambda_k \boldsymbol{d}_k)^{\mathrm{T}}\boldsymbol{d}_k \geqslant \sigma \nabla f(\boldsymbol{x}_k)^{\mathrm{T}}\boldsymbol{d}_k \\ &= \sigma\varphi'(0) > \varphi'(0).\end{aligned} \tag{6.4.4a}$$

在几何上,它表示可接受点处的切线的斜率 $\varphi'(\lambda_k)$ 应大于或等于初始斜率的 σ 倍. 我们把(6.4.2)和(6.4.4)两式称为 Wolfe - Powell 非精确线性搜索准则(Wolfe - Powell inexact line search criterion),简称为 Wolfe - Powell 准则,其可接受区间为 $J_3=[e,c]$.

算法 6 - 7(Wolfe - Powell 非精确一维搜索方法)

Step 1 选取初始数据. 在搜索区间 $[0,\infty)$(或 $[0,\lambda_{\max}]$)中取定初始点 λ_0,计算 $\varphi(0)$ 和 $\varphi'(0)$,给出可接受系数 $\rho\in(0,1)$,$\sigma\in(\rho,1)$,增大试探点系数 $\alpha>1$,令 $a_0=0, b_0=\infty$(或 $\lambda_{\max}$).

Step 2 检查下降量要求. 计算 $\varphi(\lambda_k)$,若

$$\varphi(\lambda_k) \leqslant \varphi(0) + \rho\lambda_k\varphi'(0),$$

转 Step 3;否则,令

$$a_{k+1} := a_k, b_{k+1} = \lambda_k,$$

转 Step 4.

Step 3 检查避免试探点过小要求. 计算 $\varphi'(\lambda_k)$,若

$$\varphi'(\lambda_k) \geq \sigma\varphi'(0),$$

停止迭代，输出 λ_k；否则，令

$$a_{k+1} = \lambda_k, b_{k+1} := b_k,$$

若 $b_{k+1} < \infty$，转 Step 4；否则，令

$$\lambda_{k+1} = \alpha\lambda_k, k := k + 1,$$

转 Step 2.

Step 4　选取新的试探点，取

$$\lambda_{k+1} = \frac{a_{k+1} + b_{k+1}}{2}$$

令 $k := k + 1$，转 Step 2.

在 Wolfe - Powell 法中，通常建议取 $\rho = 0.1$，$\sigma \in [0.6, 0.8]$，$\alpha = 2$.

6.4.3　简单准则与后退法

在实际中有时仅采用准则(6.4.2)式，并要求 λ 不太小. 我们把仅利用准则(6.4.2)式的做法叫做简单准则. 利用简单准则的非精确一维搜索方法称为后退法. 其基本思想是，开始令 $\lambda = 1$，如果 $\boldsymbol{x}_k + \lambda\boldsymbol{d}_k$ 不可接受，则减少 λ（后退），一直到 $\boldsymbol{x}_k + \lambda\boldsymbol{d}_k$ 可接受为止.

算法 6 - 8（后退法）

给定 $\rho \in \left(0, \frac{1}{2}\right)$，$0 < l < u < 1$.

Step 1　令 $\lambda = 1$.

Step 2　如果

$$f(\boldsymbol{x}_k + \lambda\boldsymbol{d}_k) \leq f(\boldsymbol{x}_k) + \rho\lambda\boldsymbol{g}_k^{\mathrm{T}}\boldsymbol{d}_k$$

则令 $\lambda_k = \lambda$，停止迭代，输出 λ_k；否则，转 Step 3.

Step 3　令 $\lambda := \omega\lambda$，$\omega \in [l, u]$，转 Step 2.

习　题　六

1. 设 $\bar{\boldsymbol{x}}$ 是正定二次函数 $f(\boldsymbol{x}) = \frac{1}{2}\boldsymbol{x}^{\mathrm{T}}\boldsymbol{Q}\boldsymbol{x} + \boldsymbol{b}^{\mathrm{T}}\boldsymbol{x} + c$ 的极小点，$\boldsymbol{p}$ 是 $\boldsymbol{Q}$ 的属于特征值 λ 的特征向量，$\hat{\boldsymbol{x}} = \bar{\boldsymbol{x}} + \mu\boldsymbol{p}$ $(\hat{\boldsymbol{x}} \neq \bar{\boldsymbol{x}})$，证明：

(1) $\nabla f(\hat{\boldsymbol{x}}) = \mu\lambda\boldsymbol{p}$；

(2) 从 $\hat{\boldsymbol{x}}$ 出发，沿最速下降方向作精确的一维搜索，一步达到极小点 $\bar{\boldsymbol{x}}$.

2. 设 $f(x) = \frac{1}{2}\boldsymbol{x}^{\mathrm{T}}\boldsymbol{Q}\boldsymbol{x} + \boldsymbol{b}^{\mathrm{T}}\boldsymbol{x} + c$ 是正定二次函数，设从点 $\boldsymbol{x}_k$ 处沿搜索方向 $\boldsymbol{d}_k$ 作一维精确线性搜索得后继迭代点 $\boldsymbol{x}_{k+1}$，证明下述迭代公式

$$\boldsymbol{x}_{k+1} = \boldsymbol{x}_k - \frac{\boldsymbol{d}_k^{\mathrm{T}} \nabla f(\boldsymbol{x}_k)}{\boldsymbol{d}_k^{\mathrm{T}}\boldsymbol{Q}\boldsymbol{d}_k}\boldsymbol{d}_k.$$

3. 利用进退法求解

$$\min 2\lambda^2 - 3\lambda,$$

取 $\lambda_0 = 0, h = 1$，迭代 6 次.

4. 求解 $\min \varphi(t) = e^{-t} + t^2$.

（1）用 0.618 法，要求缩短后的区间长度不超过 0.2，初始区间取[0,1].

（2）用 Newton 切线法迭代两次，取初始点 $t_0 = 0$.

5. 用二次插值法求解

$$\min_{t \geq 0} \varphi(t) = t^3 - 2t + 1,$$

取初始搜索区间[0,3]，初始插值点 $t_0 = 1$，允许误差为 0.002.

6. 证明 Newton 切线法实际是一点二次插值法，即利用一点 t_k 处的函数值 $\varphi(t_k)$、一阶导数值 $\varphi'(t_k)$ 和二阶导数值 $\varphi''(t_k)$ 构造二次插值多项式

$$\psi(t) = \alpha t^2 + \beta t + \gamma,$$

并将 $\psi(t)$ 的极小点作为新迭代点 t_{k+1}.

7. 设 $\varphi: \mathbb{R} \to \mathbb{R}$ 可微，$[a_k, b_k]$ 是极小化问题 $\min\limits_{t \geq 0} \varphi(t)$ 的搜索区间，构造二次插值多项式

$$\psi(t) = \alpha t^2 + \beta t + \gamma,$$

满足条件

$$\varphi(a_k) = \psi(a_k), \varphi(b_k) = \psi(b_k), \varphi'(a_k) = \varphi'(a_k).$$

（1）试建立一个用 $\psi(t)$ 在 $[a_k, b_k]$ 上的极小点 t_{k+1} 作为 $\varphi(t)$ 在 $[a_k, b_k]$ 上的极小点的一个近似估计的一维搜索方法（二点二次插值法），推导出 t_{k+1} 的计算公式并写出计算步骤.

（2）说明割线法也是一种二点二次插值法，其中插值多项式所满足的条件为

$$\varphi(a_k) = \psi(a_k), \varphi'(a_k) = \psi'(a_k), \varphi'(b_k) = \psi'(b_k).$$

8. 在 Wolfe - Powell 法中，给出一个采用一点二次插值法求新试探点的方案，并给出 Wolfe - Powell 法相应的计算步骤和计算框图.

9. 试用 Newton 切线法的思想给出求解非线性方程 $f(x) = 0$ 的迭代公式. 取初始点 $x_0 = 0$，用该迭代公式求方程 $e^{-2x} = x$ 的根的一个近似值.

第七章　无约束最优化的解析法

无约束最优化问题的求解方法通常称为无约束最优化方法(unconstrained optimization method). 一般来说,无约束最优化问题的求解是通过一系列一维搜索来实现的. 因此,如何选择搜索方向是无约束最优化方法的核心,且不同的搜索方向形成不同的最优化方法. 本章介绍无约束最优化方法中基本的并且主要利用目标函数的解析性质(包括一阶导数和二阶导数)来构造搜索方向的一些方法,这些方法统称为解析法(analytic method).

7.1　最速下降法

考虑无约束最优化问题

$$\min f(\boldsymbol{x}), \tag{7.1.1}$$

其中 $f:\mathbb{R}^n\to\mathbb{R}$ 具有一阶连续偏导数.

人们在处理这类问题时,总希望从某一点出发,选择一个使目标函数值下降最快的方向,以便尽快到达极小点. 由 1.2 节知道,这个方向就是该点处的负梯度方向,即最速下降方向.

对于问题(7.1.1),假设已迭代了 k 次,第 k 次迭代点为 $\boldsymbol{x}_k$,且 $\nabla f(\boldsymbol{x}_k)\neq\boldsymbol{0}$,取搜索方向

$$\boldsymbol{d}_k=-\nabla f(\boldsymbol{x}_k).$$

为使目标函数值在点 $\boldsymbol{x}_k$ 处获得最快的下降,可沿 $\boldsymbol{d}_k$ 进行一维搜索. 取步长 λ_k 为最优步长,使得

$$f(\boldsymbol{x}_k+\lambda_k\boldsymbol{d}_k)=\min_{\lambda\geqslant 0}f(\boldsymbol{x}_k+\lambda\boldsymbol{d}_k),$$

得到第 $k+1$ 次迭代点

$$\boldsymbol{x}_{k+1}=\boldsymbol{x}_k+\lambda_k\boldsymbol{d}_k.$$

于是,得到点列 $\boldsymbol{x}_0,\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots$,其中 $\boldsymbol{x}_0$ 为初始点. 如果 $\nabla f(\boldsymbol{x}_k)=\boldsymbol{0}$,则 $\boldsymbol{x}_k$ 是 f 的平稳点,这时可终止迭代.

由于这种方法的每一次迭代都是沿着最速下降方向进行搜索,因此称作最速下降法(steepest descent method).

算法 7－1(最速下降法)

Step 1　选取初始数据. 选取初始点 $\boldsymbol{x}_0$，给定允许误差 $\varepsilon>0$，令 $k=0$.

Step 2　检查是否满足终止准则. 计算 $\nabla f(\boldsymbol{x}_k)$，若 $\|\nabla f(\boldsymbol{x}_k)\|<\varepsilon$，迭代终止，$\boldsymbol{x}_k$ 为问题(7.1.1)的近似最优解；否则，转 Step 3.

Step 3　进行一维搜索. 取 $\boldsymbol{d}_k=-\nabla f(\boldsymbol{x}_k)$，求 λ_k 和 $\boldsymbol{x}_{k+1}$，使得

$$f(\boldsymbol{x}_k+\lambda_k\boldsymbol{d}_k)=\min_{\lambda\geqslant 0} f(\boldsymbol{x}_k+\lambda\boldsymbol{d}_k),$$

$$\boldsymbol{x}_{k+1}=\boldsymbol{x}_k+\lambda_k\boldsymbol{d}_k.$$

令 $k:=k+1$，返回 Step 2.

如果将最速下降法应用于正定二次函数的无约束最优化问题

$$\min f(\boldsymbol{x})=\frac{1}{2}\boldsymbol{x}^{\mathrm{T}}\boldsymbol{Q}\boldsymbol{x}+\boldsymbol{b}^{\mathrm{T}}\boldsymbol{x}+c, \tag{7.1.2}$$

其中 $\boldsymbol{x}\in\mathbb{R}^n$，$\boldsymbol{Q}\in\mathbb{R}^{n\times n}$为正定矩阵，$\boldsymbol{b}\in\mathbb{R}^n$，$c\in\mathbb{R}$，则可以推出显式迭代公式.

设第 k 次迭代点为 $\boldsymbol{x}_k$，从点 $\boldsymbol{x}_k$ 出发沿 $-\nabla f(\boldsymbol{x}_k)$作一维搜索，得

$$\boldsymbol{x}_{k+1}=\boldsymbol{x}_k-\lambda_k\nabla f(\boldsymbol{x}_k),$$

其中 λ_k 为最优步长，根据定理 6.1.1，有

$$\nabla f(\boldsymbol{x}_{k+1})^{\mathrm{T}}\nabla f(\boldsymbol{x}_k)=0.$$

又因为对于问题(7.1.2)中的正定二次函数f，有

$$\nabla f(\boldsymbol{x})=\boldsymbol{Q}\boldsymbol{x}+\boldsymbol{b},\ \forall\,\boldsymbol{x}\in\mathbb{R}^n,$$

所以

$$\nabla f(\boldsymbol{x}_{k+1})=\nabla f(\boldsymbol{x}_k)-\lambda_k\boldsymbol{Q}\nabla f(\boldsymbol{x}_k),$$

从而

$$(\nabla f(\boldsymbol{x}_k)-\lambda_k\boldsymbol{Q}\nabla f(\boldsymbol{x}_k))^{\mathrm{T}}\nabla f(\boldsymbol{x}_k)=0.$$

而 $\boldsymbol{Q}$ 正定，即$\nabla f(\boldsymbol{x}_k)^{\mathrm{T}}\boldsymbol{Q}\nabla f(\boldsymbol{x}_k)>0$，故由上式解出

$$\lambda_k=\frac{\nabla f(\boldsymbol{x}_k)^{\mathrm{T}}\nabla f(\boldsymbol{x}_k)}{\nabla f(\boldsymbol{x}_k)^{\mathrm{T}}\boldsymbol{Q}\nabla f(\boldsymbol{x}_k)}, \tag{7.1.3}$$

于是

$$\boldsymbol{x}_{k+1}=\boldsymbol{x}_k-\frac{\nabla f(\boldsymbol{x}_k)^{\mathrm{T}}\nabla f(\boldsymbol{x}_k)}{\nabla f(\boldsymbol{x}_k)^{\mathrm{T}}\boldsymbol{Q}\nabla f(\boldsymbol{x}_k)}\nabla f(\boldsymbol{x}_k). \tag{7.1.4}$$

这是最速下降法用于问题(7.1.2)的迭代公式.

例 7.1.1　用最速下降法求解问题

$$\min f(\boldsymbol{x})=4x_1^2+x_2^2, \tag{7.1.5}$$

其中 $\boldsymbol{x}=(x_1,x_2)^{\mathrm{T}}$. 取初始点 $\boldsymbol{x}_0=(1,1)^{\mathrm{T}}$，允许误差 $\varepsilon=0.1$.

解　问题(7.1.5)中的f是正定二次函数，且

$$\boldsymbol{Q}=\begin{bmatrix}8&0\\0&2\end{bmatrix},\boldsymbol{b}=\begin{bmatrix}0\\0\end{bmatrix},c=0.$$

f在点 $\boldsymbol{x}=(x_1,x_2)^{\mathrm{T}}$ 处的梯度$\nabla f(\boldsymbol{x})=(8x_1,2x_2)^{\mathrm{T}}$.

第一次迭代：

令搜索方向 $\boldsymbol{d}_0=-\nabla f(\boldsymbol{x}_0)=(-8,-2)^{\mathrm{T}}$,

$$\|\boldsymbol{d}_0\|=\sqrt{64+4}=2\sqrt{17}>\varepsilon,$$

从点 $\boldsymbol{x}_0$ 出发沿 $\boldsymbol{d}_0$ 作一维搜索，由(7.1.3)式和(7.1.4)式有

$$\lambda_0=\frac{68}{520}=0.130\ 769,$$

$$\boldsymbol{x}_1=(1,1)^{\mathrm{T}}+0.130\ 769(-8,-2)^{\mathrm{T}}=(-0.046\ 152,0.738\ 462)^{\mathrm{T}}.$$

第二次迭代：

令 $\boldsymbol{d}_1=-\nabla f(\boldsymbol{x}_1)=(0.369\ 216,-1.476\ 924)^{\mathrm{T}}$,

$$\|\boldsymbol{d}_1\|=\sqrt{2.183\ 05}=1.522\ 375>\varepsilon,$$

从点 $\boldsymbol{x}_1$ 出发沿 $\boldsymbol{d}_1$ 作一维搜索，按(7.1.4)式得

$$\boldsymbol{x}_2=(0.101\ 537,0.147\ 682)^{\mathrm{T}}.$$

第三次迭代：

令 $\boldsymbol{d}_2=-\nabla f(\boldsymbol{x}_2)=(-0.812\ 296,-0.295\ 364)^{\mathrm{T}}$,

$$\|\boldsymbol{d}_2\|=\sqrt{0.747\ 056}=0.864\ 329>\varepsilon,$$

按(7.1.4)式求得

$$\boldsymbol{x}_3=(-0.009\ 747,0.107\ 217)^{\mathrm{T}}.$$

第四次迭代：

令 $\boldsymbol{d}_3=-\nabla f(\boldsymbol{x}_3)=(0.077\ 976,-0.214\ 434)^{\mathrm{T}}$,

$$\|\boldsymbol{d}_3\|=\sqrt{0.052\ 062}=0.228\ 171>\varepsilon,$$

按(7.1.4)式求得

$$\boldsymbol{x}_4=(0.019\ 126,0.027\ 816)^{\mathrm{T}}.$$

第五次迭代：

令 $\boldsymbol{d}_4=-\nabla f(\boldsymbol{x}_4)=(-0.153\ 008,-0.055\ 632)^{\mathrm{T}}$,

$$\|\boldsymbol{d}_4\|=\sqrt{0.026\ 506}=0.162\ 807>\varepsilon,$$

按(7.1.4)式求得

$$\boldsymbol{x}_5=(-0.001\ 835,0.020\ 195)^{\mathrm{T}}.$$

此时，$\|\nabla f(\boldsymbol{x}_5)\|=\sqrt{0.001\ 847}<\varepsilon$，已满足精度要求，故得问题(7.1.5)的近似最优解

$$\boldsymbol{x}_5=(-0.001\ 835,0.020\ 195)^{\mathrm{T}}.$$

实际上，问题(7.1.5)的最优解为 $\bar{\boldsymbol{x}}=(0,0)^{\mathrm{T}}$. □

算法的收敛性由下面的定理保证.

定理 7.1.1 设 $f:\mathbb{R}^n\to\mathbb{R}$ 具有一阶连续偏导数，$\boldsymbol{x}_0\in\mathbb{R}^n$，记 $\alpha=f(\boldsymbol{x}_0)$，假定水平集 $S(f,\alpha)$ 有界，令 $\{\boldsymbol{x}_k\}$ 是由最速下降法求解问题(7.1.1)产生的点列，则

（1）当$\{x_k\}$是有穷点列时，其最后一个点是f的平稳点；

（2）当$\{x_k\}$是无穷点列时，它必有极限点，并且任一极限点都是f的平稳点.

证明　（1）当$\{x_k\}$是有穷点列时，由最速下降法的终止准则可知，其最后一个点$\bar{x}$满足$\nabla f(\bar{x})=\mathbf{0}$，即$\bar{x}$为$f$的平稳点.

（2）当$\{x_k\}$是无穷点列时，有

$$d_k=-\nabla f(x_k)\neq\mathbf{0},k=0,1,2,\cdots,$$

从而由f在点x_k处的 Taylor 公式

$$f(x_k+\lambda d_k)=f(x_k)+\lambda\nabla f(x_k)^{\mathrm{T}}d_k+o(\lambda\|d_k\|)$$

知，对充分小的$\lambda>0$，有

$$f(x_k+\lambda d_k)=f(x_k)-\lambda\|\nabla f(x_k)\|^2+o(\lambda\|d_k\|)<f(x_k).$$

故$f(x_{k+1})=f(x_k+\lambda_k d_k)=\min\limits_{\lambda\geqslant0}f(x_k+\lambda d_k)<f(x_k)$. 因此

$$f(x_{k+1})<f(x_0)=\alpha,k=0,1,2,\cdots.$$

所以数列$\{f(x_k)\}$是单调减小的，且$\{x_k\}\subseteq S(f,\alpha)$，又因$S(f,\alpha)$为有界闭集，故连续函数$f$在$S(f,\alpha)$上有界. 于是，$\{f(x_k)\}$存在极限，记

$$\lim_{k\to\infty}f(x_k)=\bar{f}.$$

根据 Bolzano - Weierstrass 定理，有界点列$\{x_k\}$必有极限点，即$\{x_k\}$中存在收敛子列$\{x_{k_m}\}$，记

$$\lim_{m\to\infty}x_{k_m}=\bar{x}.$$

由f的连续性知

$$\bar{f}=\lim_{m\to\infty}f(x_{k_m})=f(\lim_{m\to\infty}x_{k_m})=f(\bar{x}).\qquad(7.1.6)$$

现在用反证法证明$\nabla f(\bar{x})=\mathbf{0}$. 若不然，$-\nabla f(\bar{x})\neq\mathbf{0}$，对充分小的$\lambda>0$，有

$$f(\bar{x}-\lambda\nabla f(\bar{x}))<f(\bar{x}).$$

由于

$$f(x_{k+1})=f(x_k+\lambda_k d_k)\leqslant f(x_k+\lambda d_k),k=0,1,2,\cdots,$$

因此

$$f(x_{k_m+1})\leqslant f(x_{k_m}-\lambda\nabla f(x_{k_m})),m=1,2,\cdots.$$

注意到f及其偏导数连续，故令$m\to\infty$，有

$$\bar{f}\leqslant f(\bar{x}-\lambda\nabla f(\bar{x}))<f(\bar{x}),$$

此与(7.1.6)式矛盾. 这就证明了$\nabla f(\bar{x})=\mathbf{0}$，即$\bar{x}$为$f$的平稳点. □

所谓最速下降方向$-\nabla f(x_k)$仅仅反映了f在点x_k处的局部性质，对局部来说是最速下降方向，但对整个求解过程并不一定使目标值下降得最快. 事实上，由定理 6.1.1 知，在最速下降法中相继两次迭代的搜索方向是正交的，即

$$\nabla f(x_{k+1})^{\mathrm{T}}\nabla f(x_k)=0,$$

由此可见，最速下降法逼近极小点 $\bar{x}$ 的路线是锯齿形（zigzag）的，当迭代点越靠近 $\bar{x}$，其搜索步长就越小，因而收敛速度越慢（见图 7.1.1）.

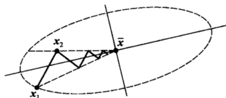


图 7.1.1　最速下降法的锯齿现象

下面我们证明最速下降法仅具有线性收敛速度.

定理 7.1.2　设 $f:\mathbb{R}^n\to\mathbb{R}$ 具有二阶连续偏导数，由最速下降法解问题(7.1.1)产生的点列 $\{x_k\}$ 收敛于 $\bar{x}$. 若存在 $\varepsilon>0$ 和 $M>m>0$，使得当 $\|x-\bar{x}\|<\varepsilon$ 时，有

$$m\|y\|^2\leqslant y^{\mathrm{T}}\nabla^2 f(x)y\leqslant M\|y\|^2,\ \forall y\in\mathbb{R}^n, \tag{7.1.7}$$

则 $\{x_k\}$ 线性收敛于 $\bar{x}$.

在证明这个定理之前，先证明一个引理.

引理 7.1.3　在定理 7.1.2 的假设条件下，当 $\|x-\bar{x}\|<\varepsilon$ 时，有

$$\frac{1}{2}m\|x-\bar{x}\|^2\leqslant f(x)-f(\bar{x})\leqslant\frac{1}{2}M\|x-\bar{x}\|^2, \tag{7.1.8}$$

$$\|\nabla f(x)\|\geqslant m\|x-\bar{x}\|. \tag{7.1.9}$$

证明　由于 f 具有二阶连续偏导数，记 $p=x-\bar{x}$，由此有

$$\begin{aligned}f(x)-f(\bar{x})&=\int_0^1 \mathrm{d}f(\bar{x}+tp)=-\int_0^1\nabla f(\bar{x}+tp)^{\mathrm{T}}p\,\mathrm{d}(1-t)\\&=-(1-t)\nabla f(\bar{x}+tp)^{\mathrm{T}}p\Big|_0^1+\int_0^1(1-t)\,\mathrm{d}[\nabla f(\bar{x}+tp)^{\mathrm{T}}p]\\&=\nabla f(\bar{x})^{\mathrm{T}}p+\int_0^1(1-t)p^{\mathrm{T}}\nabla^2 f(\bar{x}+tp)p\,\mathrm{d}t,\end{aligned}$$

由定理 7.1.1 知 $\nabla f(\bar{x})=\mathbf{0}$，故得

$$f(x)-f(\bar{x})=\int_0^1(1-t)(x-\bar{x})^{\mathrm{T}}\nabla^2 f(tx+(1-t)\bar{x})(x-\bar{x})\,\mathrm{d}t. \tag{7.1.10}$$

设 $\|x-\bar{x}\|<\varepsilon$，则由 $0\leqslant t\leqslant 1$ 易知，$\|tx+(1-t)\bar{x}-\bar{x}\|<\varepsilon$，从而由(7.1.7)式有

$$m\|x-\bar{x}\|^2\leqslant(x-\bar{x})^{\mathrm{T}}\nabla^2 f(tx+(1-t)\bar{x})(x-\bar{x})\leqslant M\|x-\bar{x}\|^2, \tag{7.1.11}$$

根据(7.1.10)(7.1.11)两式即得(7.1.8)式.

同理可得

$$\nabla f(\boldsymbol{x}) = \nabla f(\boldsymbol{x}) - \nabla f(\bar{\boldsymbol{x}}) = \int_0^1 \nabla^2 f(t\boldsymbol{x} + (1-t)\bar{\boldsymbol{x}})(\boldsymbol{x} - \bar{\boldsymbol{x}})\mathrm{d}t,$$

于是,由 Cauchy 不等式及(7.1.11)式知

$$\|\nabla f(\boldsymbol{x})\| \cdot \|\boldsymbol{x} - \bar{\boldsymbol{x}}\| \geqslant (\boldsymbol{x} - \bar{\boldsymbol{x}})^{\mathrm{T}} \nabla f(\boldsymbol{x})$$

$$= \int_0^1 (\boldsymbol{x} - \bar{\boldsymbol{x}})^{\mathrm{T}} \nabla^2 f(t\boldsymbol{x} + (1-t)\bar{\boldsymbol{x}})(\boldsymbol{x} - \bar{\boldsymbol{x}})\mathrm{d}t \geqslant m\|\boldsymbol{x} - \bar{\boldsymbol{x}}\|^2.$$

这就证明了(7.1.9)式. □

定理 7.1.2 的证明 因为 $\lim\limits_{k\to\infty}\boldsymbol{x}_k = \bar{\boldsymbol{x}}$,所以不妨假定:

$$\|\boldsymbol{x}_k - \bar{\boldsymbol{x}}\| < \varepsilon, k = 0,1,2,\cdots.$$

由于 $\|\boldsymbol{x}_{k+1} - \bar{\boldsymbol{x}}\| < \varepsilon$,因此存在 $\delta > 0$,使

$$\|\boldsymbol{x}_k + (\lambda_k + \delta)\boldsymbol{d}_k - \bar{\boldsymbol{x}}\| = \|\boldsymbol{x}_{k+1} - \bar{\boldsymbol{x}} + \delta\boldsymbol{d}_k\| < \varepsilon.$$

考虑函数 $\varphi(\lambda) = f(\boldsymbol{x}_k + \lambda\boldsymbol{d}_k)$,则由 $\boldsymbol{d}_k = -\nabla f(\boldsymbol{x}_k)$ 知:

$$\varphi'(0) = \nabla f(\boldsymbol{x}_k)^{\mathrm{T}}\boldsymbol{d}_k < 0,$$

$$\varphi'(\lambda_k) = \nabla f(\boldsymbol{x}_k + \lambda_k\boldsymbol{d}_k)^{\mathrm{T}}\boldsymbol{d}_k = 0,$$

$$\varphi''(\lambda) = \boldsymbol{d}_k^{\mathrm{T}} \nabla^2 f(\boldsymbol{x}_k + \lambda\boldsymbol{d}_k)\boldsymbol{d}_k \leqslant M\|\boldsymbol{d}_k\|^2.$$

记 $\psi(\lambda) = \varphi(0) + M\|\boldsymbol{d}_k\|^2\lambda$,它有惟一零点

$$\tilde{\lambda}_k = \frac{-\varphi'(0)}{M\|\boldsymbol{d}_k\|^2}.$$

于是

$$\varphi'(\lambda) = \varphi'(0) + \int_0^\lambda \varphi''(t)\mathrm{d}t \leqslant \varphi'(0) + \int_0^\lambda M\|\boldsymbol{d}_k\|^2\mathrm{d}t$$

$$= \varphi'(0) + M\|\boldsymbol{d}_k\|^2\lambda = \psi(\lambda).$$

上式说明,函数 $\varphi'(\lambda)$ 的图像总是在 $\psi(\lambda)$ 的图像之下(见图 7.1.2),从而 $\varphi(\lambda)$ 的平稳点 λ_k 满足

$$\lambda_k \geqslant \tilde{\lambda}_k = \frac{-\varphi'(0)}{M\|\boldsymbol{d}_k\|^2} = \frac{\boldsymbol{d}_k^{\mathrm{T}}\boldsymbol{d}_k}{M\|\boldsymbol{d}_k\|^2} = \frac{1}{M}.$$

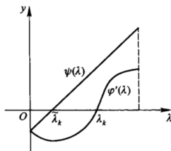


图 7.1.2 $\varphi'(\lambda)$ 与 $\psi(\lambda)$ 的关系

令 $\tilde{\boldsymbol{x}}_k = \boldsymbol{x}_k + \tilde{\lambda}_k\boldsymbol{d}_k$,显然有

$$\|\tilde{\boldsymbol{x}}_k - \bar{\boldsymbol{x}}\| \leqslant \|\boldsymbol{x}_k + \lambda_k\boldsymbol{d}_k - \bar{\boldsymbol{x}}\| < \varepsilon,$$

于是,仿照引理 7.1.3 的证明可得

$$
\begin{aligned}
f(\boldsymbol{x}_k+\lambda_k\boldsymbol{d}_k)-f(\boldsymbol{x}_k) &\leqslant f(\boldsymbol{x}_k+\tilde{\lambda}_k\boldsymbol{d}_k)-f(\boldsymbol{x}_k)\\
&=\tilde{\lambda}_k\nabla f(\boldsymbol{x}_k)^{\mathrm{T}}\boldsymbol{d}_k+\tilde{\lambda}_k^2\int_0^1(1-t)\boldsymbol{d}_k^{\mathrm{T}}\nabla^2 f(\boldsymbol{x}_k+t\tilde{\lambda}_k\boldsymbol{d}_k)\boldsymbol{d}_k\mathrm{d}t\\
&\leqslant-\tilde{\lambda}_k\|\boldsymbol{d}_k\|^2+\frac{1}{2}\tilde{\lambda}_k^2M\|\boldsymbol{d}_k\|^2\\
&=-\frac{\|\boldsymbol{d}_k\|^2}{M}+\frac{\|\boldsymbol{d}_k\|^2}{2M}\\
&=-\frac{\|\boldsymbol{d}_k\|^2}{2M},
\end{aligned}
$$

从而由(7.1.9)和(7.1.8)两式得

$$
f(\boldsymbol{x}_{k+1})-f(\boldsymbol{x}_k)\leqslant-\frac{m^2}{2M}\|\boldsymbol{x}_k-\bar{\boldsymbol{x}}\|^2\leqslant-\left(\frac{m}{M}\right)^2[f(\boldsymbol{x}_k)-f(\bar{\boldsymbol{x}})],
$$

即知

$$
f(\boldsymbol{x}_{k+1})-f(\bar{\boldsymbol{x}})\leqslant\left[1-\left(\frac{m}{M}\right)^2\right][f(\boldsymbol{x}_k)-f(\bar{\boldsymbol{x}})].
$$

令

$$
\theta=\left[1-\left(\frac{m}{M}\right)^2\right]^{\frac{1}{2}},
$$

显然 $\theta\in(0,1)$,故有

$$
\begin{aligned}
f(\boldsymbol{x}_k)-f(\bar{\boldsymbol{x}}) &\leqslant\theta^2[f(\boldsymbol{x}_{k-1})-f(\bar{\boldsymbol{x}})]\leqslant\cdots\\
&\leqslant\theta^{2k}[f(\boldsymbol{x}_0)-f(\bar{\boldsymbol{x}})],
\end{aligned}
$$

由(7.1.8)式得

$$
\|\boldsymbol{x}_k-\bar{\boldsymbol{x}}\|^2\leqslant\frac{2}{m}[f(\boldsymbol{x}_k)-f(\bar{\boldsymbol{x}})]\leqslant\frac{2}{m}[f(\boldsymbol{x}_0)-f(\bar{\boldsymbol{x}})]\theta^{2k},
$$

即

$$
\|\boldsymbol{x}_k-\bar{\boldsymbol{x}}\|\leqslant\sqrt{\frac{2}{m}}[f(\boldsymbol{x}_0)-f(\bar{\boldsymbol{x}})]^{1/2}\theta^k,
$$

其中 $\theta<1$. 这就证明了$\{\boldsymbol{x}_k\}$线性收敛于 $\bar{\boldsymbol{x}}$. □

7.2 Newton 法

为了寻找收敛速度快的无约束最优化方法,我们考虑在每次迭代时,用适当的二次函数去近似目标函数f,并用迭代点指向近似二次函数极小点的方向来构造搜索方向,然后精确地求出近似二次函数的极小点,以该极小点作为f的极小点的近似值. 这就是 Newton 法的基本思想,它是 6.3.1 节中 Newton 切线法的推广.

假定问题(7.1.1)中的目标函数 f 具有二阶连续偏导数，$\boldsymbol{x}_k$ 是 f 的极小点的第 k 次近似，将 f 在点 $\boldsymbol{x}_k$ 处作 Taylor 展开，并取二阶近似，得

$$f(\boldsymbol{x}) \approx \varphi(\boldsymbol{x}) = f(\boldsymbol{x}_k) + \nabla f(\boldsymbol{x}_k)^{\mathrm{T}}(\boldsymbol{x}-\boldsymbol{x}_k) + \frac{1}{2}(\boldsymbol{x}-\boldsymbol{x}_k)^{\mathrm{T}} \nabla^2 f(\boldsymbol{x}_k)(\boldsymbol{x}-\boldsymbol{x}_k).$$

由假设条件知，$\nabla^2 f(\boldsymbol{x}_k)$ 是对称矩阵，因此 $\varphi(\boldsymbol{x})$ 是二次函数. 为求 $\varphi(\boldsymbol{x})$ 的极小点，可令 $\nabla \varphi(\boldsymbol{x}) = \boldsymbol{0}$，即

$$\nabla f(\boldsymbol{x}_k) + \nabla^2 f(\boldsymbol{x}_k)(\boldsymbol{x}-\boldsymbol{x}_k) = \boldsymbol{0}.$$

若 f 在点 $\boldsymbol{x}_k$ 处的 Hesse 矩阵 $\nabla^2 f(\boldsymbol{x}_k)$ 正定，则上式解出的 $\varphi(\boldsymbol{x})$ 的平稳点就是 $\varphi(\boldsymbol{x})$ 的极小点. 以它作为 f 的极小点的第 $k+1$ 次近似，记为 $\boldsymbol{x}_{k+1}$，即

$$\boldsymbol{x}_{k+1} = \boldsymbol{x}_k - [\nabla^2 f(\boldsymbol{x}_k)]^{-1} \nabla f(\boldsymbol{x}_k). \tag{7.2.1}$$

这就是 Newton 法的迭代公式，其中

$$\boldsymbol{d}_k = -[\nabla^2 f(\boldsymbol{x}_k)]^{-1} \nabla f(\boldsymbol{x}_k)$$

称为 Newton 方向. 它是第 $k+1$ 次迭代的搜索方向，且步长为 1.

因 $\nabla^2 f(\boldsymbol{x}_k)$ 正定，故 $[\nabla^2 f(\boldsymbol{x}_k)]^{-1}$ 正定，从而

$$\nabla f(\boldsymbol{x}_k)^{\mathrm{T}} \boldsymbol{d}_k = -\nabla f(\boldsymbol{x}_k)^{\mathrm{T}} [\nabla^2 f(\boldsymbol{x}_k)]^{-1} \nabla f(\boldsymbol{x}_k) < 0,$$

所以由定理 1.2.3 知，$\boldsymbol{d}_k$ 为 f 在点 $\boldsymbol{x}_k$ 处的下降方向.

算法 7 - 2(Newton 法)

Step 1　选取初始数据. 选取初始点 $\boldsymbol{x}_0$，给定允许误差 $\varepsilon > 0$，令 $k = 0$.

Step 2　检验是否满足终止准则. 计算 $\nabla f(\boldsymbol{x}_k)$，若 $\|\nabla f(\boldsymbol{x}_k)\| < \varepsilon$，迭代终止，$\boldsymbol{x}_k$ 为问题(7.1.1)的近似最优解；否则，转 Step 3.

Step 3　构造 Newton 方向. 计算 $[\nabla^2 f(\boldsymbol{x}_k)]^{-1}$，取

$$\boldsymbol{d}_k = -[\nabla^2 f(\boldsymbol{x}_k)]^{-1} \nabla f(\boldsymbol{x}_k).$$

Step 4　求下一个迭代点. 令

$$\boldsymbol{x}_{k+1} = \boldsymbol{x}_k + \boldsymbol{d}_k, k := k+1,$$

返回 Step 2.

我们知道，正定二次函数的无约束最优化问题(7.1.2)的全局极小点为 $\bar{\boldsymbol{x}} = -Q^{-1}\boldsymbol{b}$. 因此，如果对问题(7.1.2)用 Newton 法迭代，从任一点 $\boldsymbol{x}_0 \in \mathbb{R}^n$ 出发，可得

$$\boldsymbol{x}_1 = \boldsymbol{x}_0 - Q^{-1} \nabla f(\boldsymbol{x}_0) = \boldsymbol{x}_0 - Q^{-1}(Q\boldsymbol{x}_0 + \boldsymbol{b}) = -Q^{-1}\boldsymbol{b} = \bar{\boldsymbol{x}},$$

即一次迭代就可得到全局极小点. 这说明，Newton 法具有二次终止性.

例 7.2.1　用 Newton 法求解问题(7.1.5)，仍取初始点 $\boldsymbol{x}_0 = (1,1)^{\mathrm{T}}$，允许误差 $\varepsilon = 0.1$.

解　$\nabla f(\boldsymbol{x}_0) = (8,2)^{\mathrm{T}}, \nabla^2 f(\boldsymbol{x}_0) = \begin{bmatrix} 8 & 0 \\ 0 & 2 \end{bmatrix}$，故

$$[\nabla^2 f(\boldsymbol{x}_0)]^{-1} = \begin{bmatrix} 1/8 & 0 \\ 0 & 1/2 \end{bmatrix}, \boldsymbol{d}_0 = -[\nabla^2 f(\boldsymbol{x}_0)]^{-1}\nabla f(\boldsymbol{x}_0) = -\begin{bmatrix} 1 \\ 1 \end{bmatrix},$$

$$\boldsymbol{x}_1 = \boldsymbol{x}_0 + \boldsymbol{d}_0 = (1,1)^{\mathrm{T}} - (1,1)^{\mathrm{T}} = (0,0)^{\mathrm{T}}.$$

因 $\|\nabla f(\boldsymbol{x}_1)\| = 0 < 0.1$，故迭代结束，得 $\boldsymbol{x}_1$ 为问题(7.1.5)的最优解. □

这个例子说明，Newton 法要比最速下降法收敛快. 事实上，下述定理表明 Newton 法是二阶收敛的，该定理是定理 6.3.1 的一个推广.

定理 7.2.1 设 $f:\mathbb{R}^n \to \mathbb{R}$ 具有三阶连续偏导数，$\bar{\boldsymbol{x}} \in \mathbb{R}^n$，$\nabla f(\bar{\boldsymbol{x}}) = \boldsymbol{0}$，若存在 $\varepsilon > 0$ 和 $m > 0$，使得当 $\|\boldsymbol{x} - \bar{\boldsymbol{x}}\| \le \varepsilon$ 时，有

$$m\|\boldsymbol{y}\|^2 \le \boldsymbol{y}^{\mathrm{T}} \nabla^2 f(\boldsymbol{x})\boldsymbol{y}, \forall \boldsymbol{y} \in \mathbb{R}^n, \tag{7.2.2}$$

则当初始点 $\boldsymbol{x}_0$ 充分接近 $\bar{\boldsymbol{x}}$ 时，由 Newton 法解问题(7.1.1)产生的点列 $\{\boldsymbol{x}_k\}$ 收敛于 $\bar{\boldsymbol{x}}$，并有二阶收敛速度.

证明 由(7.2.2)式知，当 $\|\boldsymbol{x} - \bar{\boldsymbol{x}}\| \le \varepsilon$ 时，$\nabla^2 f(\boldsymbol{x})$ 为正定矩阵，且对一切 $\boldsymbol{y} \in \mathbb{R}^n$，有

$$\begin{aligned}\|[\nabla^2 f(\boldsymbol{x})]^{-1}\boldsymbol{y}\|^2 &\le \frac{1}{m}\boldsymbol{y}^{\mathrm{T}}[\nabla^2 f(\boldsymbol{x})]^{-1}\nabla^2 f(\boldsymbol{x})[\nabla^2 f(\boldsymbol{x})]^{-1}\boldsymbol{y} \\ &= \frac{1}{m}\boldsymbol{y}^{\mathrm{T}}[\nabla^2 f(\boldsymbol{x})]^{-1}\boldsymbol{y} \le \frac{1}{m}\|\boldsymbol{y}\|\ \|[\nabla^2 f(\boldsymbol{x})]^{-1}\boldsymbol{y}\|,\end{aligned}$$

即当 $\|\boldsymbol{x} - \bar{\boldsymbol{x}}\| \le \varepsilon$ 时，有

$$\|[\nabla^2 f(\boldsymbol{x})]^{-1}\boldsymbol{y}\| \le \frac{1}{m}\|\boldsymbol{y}\|, \forall \boldsymbol{y} \in \mathbb{R}^n.$$

由 $\nabla f(\boldsymbol{x})$ 的连续性及 $\nabla f(\bar{\boldsymbol{x}}) = \boldsymbol{0}$ 知，必存在 $\varepsilon' \in (0, \varepsilon/2)$，使得 $\|\boldsymbol{x} - \bar{\boldsymbol{x}}\| \le \varepsilon'$ 时，有

$$\|\nabla f(\boldsymbol{x})\| \le \frac{m\varepsilon}{2}.$$

这意味着：当 $\|\boldsymbol{x}_k - \bar{\boldsymbol{x}}\| \le \varepsilon'$ 时，有

$$\begin{aligned}\|\boldsymbol{x}_{k+1} - \bar{\boldsymbol{x}}\| &= \|\boldsymbol{x}_k - [\nabla^2 f(\boldsymbol{x}_k)]^{-1}\nabla f(\boldsymbol{x}_k) - \bar{\boldsymbol{x}}\| \\ &\le \|\boldsymbol{x}_k - \bar{\boldsymbol{x}}\| + \|[\nabla^2 f(\boldsymbol{x}_k)]^{-1}\nabla f(x_k)\| \\ &\le \frac{\varepsilon}{2} + \frac{1}{m}\cdot\frac{m\varepsilon}{2} = \varepsilon.\end{aligned}$$

因 $\boldsymbol{x}_0$ 充分接近 $\bar{\boldsymbol{x}}$，故可设

$$\|\boldsymbol{x}_k - \bar{\boldsymbol{x}}\| \le \varepsilon', k = 0,1,2,\cdots.$$

我们知道

$$\begin{aligned}\|\boldsymbol{x}_{k+1} - \bar{\boldsymbol{x}}\| &= \|[\nabla^2 f(\boldsymbol{x}_k)]^{-1}[\nabla f(\bar{\boldsymbol{x}}) - \nabla f(\boldsymbol{x}_k) - \nabla^2 f(\boldsymbol{x}_k)(\bar{\boldsymbol{x}} - \boldsymbol{x}_k)]\| \\ &\le \frac{1}{m}\|\nabla f(\bar{\boldsymbol{x}}) - \nabla f(\boldsymbol{x}_k) - \nabla^2 f(\boldsymbol{x}_k)(\bar{\boldsymbol{x}} - \boldsymbol{x}_k)\|.\end{aligned} \tag{7.2.3}$$

考虑向量函数 $\boldsymbol{\varphi}(\boldsymbol{x}) = \nabla f(\boldsymbol{x})$ 的第 l 个分量 $\varphi_l(\boldsymbol{x})$ 在点 $\boldsymbol{x}_k$ 处的 Taylor 展开式

$$\varphi_l(\boldsymbol{x}) = \varphi_l(\boldsymbol{x}_k) + \nabla \varphi_l(\boldsymbol{x}_k)^{\mathrm{T}}(\boldsymbol{x}-\boldsymbol{x}_k) + \frac{1}{2}(\boldsymbol{x}-\boldsymbol{x}_k)^{\mathrm{T}}\nabla^2\varphi_l(\hat{\boldsymbol{x}})(\boldsymbol{x}-\boldsymbol{x}_k),$$

其中 $\hat{\boldsymbol{x}} = \boldsymbol{x}_k + \theta(\boldsymbol{x}-\boldsymbol{x}_k)$, $0<\theta<1$. 于是

$$\nabla f(\bar{\boldsymbol{x}}) - \nabla f(\boldsymbol{x}_k) - \nabla^2 f(\boldsymbol{x}_k)(\bar{\boldsymbol{x}}-\boldsymbol{x}_k) = \frac{1}{2}\begin{bmatrix} (\bar{\boldsymbol{x}}-\boldsymbol{x}_k)^{\mathrm{T}}\nabla^2\varphi_l(\hat{\boldsymbol{x}})(\bar{\boldsymbol{x}}-\boldsymbol{x}_k) \\ (\bar{\boldsymbol{x}}-\boldsymbol{x}_k)^{\mathrm{T}}\nabla^2\varphi_{l+1}(\hat{\boldsymbol{x}})(\bar{\boldsymbol{x}}-\boldsymbol{x}_k) \\ \vdots \\ (\bar{\boldsymbol{x}}-\boldsymbol{x}_k)^{\mathrm{T}}\nabla^2\varphi_n(\hat{\boldsymbol{x}})(\bar{\boldsymbol{x}}-\boldsymbol{x}_k) \end{bmatrix}, \tag{7.2.4}$$

注意到$\nabla^2\varphi_l(\hat{\boldsymbol{x}})$的第 i 行第 j 列元素为

$$\frac{\partial^2\varphi_l(\hat{\boldsymbol{x}})}{\partial x_i \partial x_j} = \frac{\partial^3 f(\hat{\boldsymbol{x}})}{\partial x_i \partial x_j \partial x_l}.$$

因为f具有三阶连续偏导数,所以f的三阶偏导数在 $\|\boldsymbol{x}-\bar{\boldsymbol{x}}\| \leqslant \varepsilon$ 上有界,从而存在$\beta>0$,使得当 $\|\boldsymbol{x}-\bar{\boldsymbol{x}}\| \leqslant \varepsilon$ 时,有

$$\|\nabla^2\varphi_l(\boldsymbol{x})\| \leqslant \beta, l=1,2,\cdots,n.$$

而 $\hat{\boldsymbol{x}} = \boldsymbol{x}_k + \theta(\bar{\boldsymbol{x}}-\boldsymbol{x}_k)$ 且 $\|\boldsymbol{x}_k-\bar{\boldsymbol{x}}\| \leqslant \varepsilon'$,故 $\|\hat{\boldsymbol{x}}-\bar{\boldsymbol{x}}\| \leqslant \varepsilon$,易知

$$\|\nabla^2\varphi_l(\hat{\boldsymbol{x}})\| \leqslant \beta, l=1,2,\cdots,n.$$

于是,由(7.2.4)式有

$$\begin{aligned} &\|\nabla f(\bar{\boldsymbol{x}}) - \nabla f(\boldsymbol{x}_k) - \nabla^2 f(\boldsymbol{x}_k)(\bar{\boldsymbol{x}}-\boldsymbol{x}_k)\|^2 \\ &= \frac{1}{4}\sum_{i=1}^{n}[(\bar{\boldsymbol{x}}-\boldsymbol{x}_k)^{\mathrm{T}}\nabla^2\varphi_l(\hat{\boldsymbol{x}})(\bar{\boldsymbol{x}}-\boldsymbol{x}_k)]^2 \\ &\leqslant \frac{1}{4}n\beta^2\|\boldsymbol{x}_k-\bar{\boldsymbol{x}}\|^4, \end{aligned}$$

因此,由(7.2.3)式知

$$\|\boldsymbol{x}_{k+1}-\bar{\boldsymbol{x}}\| \leqslant \frac{\sqrt{n}\beta}{2m}\|\boldsymbol{x}_k-\bar{\boldsymbol{x}}\|^2, k=0,1,2,\cdots. \tag{7.2.5}$$

记 $\gamma = \dfrac{\sqrt{n}\beta}{2m}>0$,则

$$\begin{aligned} \|\boldsymbol{x}_k-\bar{\boldsymbol{x}}\| &\leqslant \gamma\|\boldsymbol{x}_{k-1}-\bar{\boldsymbol{x}}\|^2 \leqslant \cdots \\ &\leqslant \gamma^{2^k-1}\|\boldsymbol{x}_0-\bar{\boldsymbol{x}}\|^{2^k} \leqslant \frac{1}{\gamma}(\gamma\|\boldsymbol{x}_0-\bar{\boldsymbol{x}}\|)^{2^k}, \end{aligned}$$

所以,取 $0<\delta<\dfrac{1}{\gamma}$,当 $\|\boldsymbol{x}_0-\bar{\boldsymbol{x}}\|<\delta$ 时,$\{\boldsymbol{x}_k\}$收敛于 $\bar{\boldsymbol{x}}$,并且由(7.2.5)式知其具有二阶收敛速度. □

由定理 7.2.1 可知,当初始点 $\boldsymbol{x}_0$ 靠近极小点 $\bar{\boldsymbol{x}}$ 时,Newton 法的收敛速度是很快的. 但是,当 $\boldsymbol{x}_0$ 远离 $\bar{\boldsymbol{x}}$ 时,Newton 法可能不收敛,甚至连下降性也保证不了.

其原因是：迭代点 $\boldsymbol{x}_{k+1}$ 不一定是目标函数 f 在 Newton 方向 $\boldsymbol{d}_k$ 上的极小点. 下面的例子就说明了这一点.

例 7.2.2 用 Newton 法求解下面问题：

$$\min f(\boldsymbol{x}) = (1-x_1)^2 + 2(x_2 - x_1^2)^2, \tag{7.2.6}$$

其中 $\boldsymbol{x} = (x_1, x_2)^{\mathrm{T}}$，取初始点 $\boldsymbol{x}_0 = (0,0)^{\mathrm{T}}$，允许误差 $\varepsilon = 0.1$.

解

$$\nabla f(\boldsymbol{x}) = \begin{bmatrix} -2(1-x_1) - 8(x_2 - x_1^2)x_1 \\ 4(x_2 - x_1^2) \end{bmatrix},$$

$$\nabla^2 f(\boldsymbol{x}) = \begin{bmatrix} 16x_1^2 - 8(x_2 - x_1^2) + 2 & -8x_1 \\ -8x_1 & 4 \end{bmatrix},$$

故

$$\nabla f(\boldsymbol{x}_0) = (-2,0)^{\mathrm{T}}, \ \|\nabla f(\boldsymbol{x}_0)\| = 2 > \varepsilon,$$

$$\nabla^2 f(\boldsymbol{x}_0) = \begin{bmatrix} 2 & 0 \\ 0 & 4 \end{bmatrix}, [\nabla^2 f(\boldsymbol{x}_0)]^{-1} = \begin{bmatrix} 1/2 & 0 \\ 0 & 1/4 \end{bmatrix}.$$

于是，Newton 方向

$$\boldsymbol{d}_0 = -[\nabla^2 f(\boldsymbol{x}_0)]^{-1} \nabla f(\boldsymbol{x}_0) = (1,0)^{\mathrm{T}},$$

$$\boldsymbol{x}_1 = \boldsymbol{x}_0 + \boldsymbol{d}_0 = (1,0)^{\mathrm{T}}.$$

但 $f(\boldsymbol{x}_0) = 1, f(\boldsymbol{x}_1) = 2 > f(\boldsymbol{x}_0)$，即知迭代后未使目标函数值下降，反而使目标函数值上升. □

为了弥补 Newton 法的上述缺陷，人们把 Newton 法作了如下修正：由 $\boldsymbol{x}_k$ 求 $\boldsymbol{x}_{k+1}$ 时，不直接用迭代公式(7.2.1)而是沿 Newton 方向 $\boldsymbol{d}_k$ 进行最优一维搜索. 这样就是所谓的阻尼 Newton 法(damped Newton method).

算法 7-3(阻尼 Newton 法)

Step 1 选取初始数据. 选取初始点 $\boldsymbol{x}_0$，给定允许误差 $\varepsilon > 0$，令 $k = 0$.

Step 2 检验是否满足终止准则. 计算 $\nabla f(\boldsymbol{x}_k)$，若 $\|\nabla f(\boldsymbol{x}_k)\| < \varepsilon$，迭代终止，$\boldsymbol{x}_k$ 为问题(7.1.1)的近似最优解；否则，转 Step 3.

Step 3 构造 Newton 方向. 计算 $[\nabla^2 f(\boldsymbol{x}_k)]^{-1}$，取

$$\boldsymbol{d}_k = -[\nabla^2 f(\boldsymbol{x}_k)]^{-1} \nabla f(\boldsymbol{x}_k).$$

Step 4 进行一维搜索. 求 λ_k 和 $\boldsymbol{x}_{k+1}$，使得

$$f(\boldsymbol{x}_k + \lambda_k \boldsymbol{d}_k) = \min_{\lambda \geqslant 0} f(\boldsymbol{x}_k + \lambda \boldsymbol{d}_k),$$

$$\boldsymbol{x}_{k+1} = \boldsymbol{x}_k + \lambda_k \boldsymbol{d}_k,$$

令 $k := k+1$，返回 Step 2.

下面的定理指出，阻尼 Newton 法具有全局收敛性.

定理 7.2.2 设 $f: \mathbb{R}^n \to \mathbb{R}$ 具有二阶连续偏导数，$\boldsymbol{x}_0 \in \mathbb{R}^n$，记 $\alpha = f(\boldsymbol{x}_0)$，假定水平集 $S(f,\alpha)$ 有界，并且对一切 $\boldsymbol{x} \in \mathbb{R}^n$，$\nabla^2 f(\boldsymbol{x})$ 正定. 若 $\{\boldsymbol{x}_k\}$ 是由阻尼 Newton 法求解问题(7.1.1)产生的点列，则

（1）当$\{x_k\}$是有穷点列时，其最后一个点必是f的惟一极小点；

（2）当$\{x_k\}$是无穷点列时，它必收敛于f的惟一极小点.

证明 因为f具有二阶连续偏导数，且对一切$x \in \mathbb{R}^n$，$\nabla^2 f(x)$正定，所以由定理3.1.3知，f的平稳点必定是f的惟一全局极小点.

仿照定理7.1.1的证明可知，总存在$\bar{x} \in \mathbb{R}^n$，使$\nabla f(\bar{x}) = \mathbf{0}$，由前知，$\bar{x}$为$f$的惟一极小点. 从而定理得证. □

例 7.2.3 用阻尼Newton法求解问题(7.2.6)，仍取初始点$x_0 = (0,0)^T$，允许误差$\varepsilon = 0.1$.

解 第一次迭代：

由例7.2.2知，Newton方向$d_0 = (1,0)^T$，从x_0出发沿d_0作一维搜索，即求

$$\min_{\lambda \geqslant 0} f(x_0 + \lambda d_0) = \min_{\lambda \geqslant 0} (1 - \lambda)^2 + 2\lambda^4$$

的最优解，得到$\lambda_0 = 1/2$. 令

$$x_1 = x_0 + \lambda_0 d_0 = (1/2, 0)^T,$$

$$\nabla f(x_1) = (0, -1)^T,\ \|\nabla f(x_1)\| = 1 > \varepsilon.$$

第二次迭代：

$$\nabla^2 f(x_1) = \begin{bmatrix} 8 & -4 \\ -4 & 4 \end{bmatrix},\ [\nabla^2 f(x_1)]^{-1} = \frac{1}{4}\begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix},$$

$$d_1 = -[\nabla^2 f(x_1)]^{-1} \nabla f(x_1) = (1/4, 1/2)^T.$$

从x_1出发，沿d_1作一维搜索，即求

$$\min_{\lambda \geqslant 0} f(x_1 + \lambda d_1) = \min_{\lambda \geqslant 0} \frac{1}{128}[8(2-\lambda)^2 + (2-\lambda)^4]$$

的最优解，得到$\lambda_1 = 2$. 令

$$x_2 = x_1 + \lambda_1 d_1 = (1,1)^T.$$

此时

$$\nabla f(x_2) = (0,0)^T,\ \|\nabla f(x_2)\| = 0 < \varepsilon.$$

得问题(7.2.6)的最优解为$x_2 = (1,1)^T$，这是惟一的最优解. □

7.3 共轭梯度法

最速下降法和Newton法是最基本的无约束最优化方法，它们的特性各异：前者计算量较小而收敛速度慢；后者虽然收敛速度快，但需要计算目标函数的Hesse矩阵及其逆矩阵，故计算量大. 本节介绍一类无需计算二阶导数并且收敛速度快的方法.

7.3.1 共轭方向法

定义 7.3.1 设 $\boldsymbol{Q}\in\mathbb{R}^{n\times n}$ 为正定矩阵. 若 $\mathbb{R}^n$ 中的向量组 $\boldsymbol{d}_0,\boldsymbol{d}_1,\cdots,\boldsymbol{d}_{m-1}$ 满足

$$\boldsymbol{d}_i^{\mathrm{T}}\boldsymbol{Q}\boldsymbol{d}_j=0,\ \forall\, i,j=0,1,\cdots,m-1,i\neq j,\tag{7.3.1}$$

则称 $\boldsymbol{d}_0,\boldsymbol{d}_1,\cdots,\boldsymbol{d}_{m-1}$ 是 $\boldsymbol{Q}$ 共轭的(conjugate).

当 $\boldsymbol{Q}=\boldsymbol{I}_n$ 为单位矩阵时,则(7.3.1)式变为

$$\boldsymbol{d}_i^{\mathrm{T}}\boldsymbol{d}_j=0,\ \forall\, i,j=0,1,\cdots,m-1,i\neq j,$$

即向量组 $\boldsymbol{d}_0,\boldsymbol{d}_1,\cdots,\boldsymbol{d}_{m-1}$ 是正交的(orthogonal).

由此可知,共轭是正交概念的推广.

定理 7.3.1 设 $\boldsymbol{Q}\in\mathbb{R}^{n\times n}$ 是正定矩阵,$\mathbb{R}^n$ 中非零向量组 $\boldsymbol{d}_0,\boldsymbol{d}_1,\cdots,\boldsymbol{d}_{m-1}$ 是 $\boldsymbol{Q}$ 共轭的,则这 m 个向量线性无关.

证明 若存在实数 $t_0,t_1,\cdots,t_{m-1}$,使

$$\sum_{j=0}^{m-1}t_j\boldsymbol{d}_j=\boldsymbol{0},$$

用 $\boldsymbol{d}_i^{\mathrm{T}}\boldsymbol{Q}$ 左乘上式并注意到

$$\boldsymbol{d}_i^{\mathrm{T}}\boldsymbol{Q}\boldsymbol{d}_j=0,\ \forall\, i,j=0,1,\cdots,m-1,i\neq j,$$

得到

$$t_i\boldsymbol{d}_i^{\mathrm{T}}\boldsymbol{Q}\boldsymbol{d}_i=0,i=0,1,\cdots,m-1.\tag{7.3.2}$$

但 $\boldsymbol{Q}$ 是正定矩阵,且 $\boldsymbol{d}_i\neq\boldsymbol{0},i=0,1,\cdots,m-1$,故

$$\boldsymbol{d}_i^{\mathrm{T}}\boldsymbol{Q}\boldsymbol{d}_i>0,i=0,1,\cdots,m-1.$$

从而,由(7.3.2)式知:$t_i=0,i=0,1,\cdots,m-1$,即 $\boldsymbol{d}_0,\boldsymbol{d}_1,\cdots,\boldsymbol{d}_{m-1}$ 线性无关. □

定理 7.3.2 设 $\boldsymbol{p}\in\mathbb{R}^n$,$\boldsymbol{d}_0,\boldsymbol{d}_1,\cdots,\boldsymbol{d}_{n-1}$ 是 $\mathbb{R}^n$ 中线性无关的向量组,若 $\boldsymbol{p}$ 与每个 $\boldsymbol{d}_i$ 都正交,则 $\boldsymbol{p}=\boldsymbol{0}$.

证明 因为 $\boldsymbol{d}_0,\boldsymbol{d}_1,\cdots,\boldsymbol{d}_{n-1}\in\mathbb{R}^n$ 且线性无关,所以它们构成 $\mathbb{R}^n$ 的一组基,从而 $\boldsymbol{p}\in\mathbb{R}^n$ 可以表示为它们的线性组合,即存在一组实数 $t_0,t_1,\cdots,t_{n-1}$,使

$$\boldsymbol{p}=\sum_{i=0}^{n-1}t_i\boldsymbol{d}_i.$$

由 $\boldsymbol{p}^{\mathrm{T}}\boldsymbol{p}=\boldsymbol{p}^{\mathrm{T}}\sum_{i=0}^{n-1}t_i\boldsymbol{d}_i=\sum_{i=0}^{n-1}t_i\boldsymbol{p}^{\mathrm{T}}\boldsymbol{d}_i=0$,即知 $\boldsymbol{p}=\boldsymbol{0}$. □

考虑正定二次函数的无约束最优化问题

$$\min f(\boldsymbol{x})=\frac{1}{2}\boldsymbol{x}^{\mathrm{T}}\boldsymbol{Q}\boldsymbol{x}+\boldsymbol{b}^{\mathrm{T}}\boldsymbol{x}+c,\tag{7.3.3}$$

其中 $\boldsymbol{Q}\in\mathbb{R}^{n\times n}$ 为正定矩阵,$\boldsymbol{b}\in\mathbb{R}^n$,$c\in\mathbb{R}$.

问题(7.3.3)有惟一的严格全局最优解 $\bar{\boldsymbol{x}}=-\boldsymbol{Q}^{-1}\boldsymbol{b}$,下面我们不求 $\boldsymbol{Q}^{-1}$,而

是用迭代的方法求问题(7.3.3)的最优解.

当 $n=2$ 时,任选初始点 $\boldsymbol{x}_0$,沿 f 在点 $\boldsymbol{x}_0$ 处的某个下降方向 $\boldsymbol{d}_0$ 做最优一维搜索,得到 $\boldsymbol{x}_1$,从而

$$\nabla f(\boldsymbol{x}_1)^{\mathrm{T}}\boldsymbol{d}_0=0. \tag{7.3.4}$$

如果再沿最速下降方向 $-\nabla f(\boldsymbol{x}_1)$ 去搜索,就会发生锯齿现象. 为了避免这种现象的出现,我们希望下一次迭代的搜索方向 $\boldsymbol{d}_1$ 直指正定二次函数 f 的最优点 $\bar{\boldsymbol{x}}$,即有

$$\bar{\boldsymbol{x}}=\boldsymbol{x}_1+\lambda_1\boldsymbol{d}_1. \tag{7.3.5}$$

显然,当 $\bar{\boldsymbol{x}}\neq\boldsymbol{x}_1$ 时,$\lambda_1\neq0$,由 $\bar{\boldsymbol{x}}$ 为 f 的极小点知

$$\boldsymbol{0}=\nabla f(\bar{\boldsymbol{x}})=\boldsymbol{Q}\bar{\boldsymbol{x}}+\boldsymbol{b}.$$

把(7.3.5)代入上式,并注意到 $\nabla f(\boldsymbol{x}_1)=\boldsymbol{Q}\boldsymbol{x}_1+\boldsymbol{b}$,得

$$\nabla f(\boldsymbol{x}_1)+\lambda_1\boldsymbol{Q}\boldsymbol{d}_1=\boldsymbol{0}.$$

上式两端左乘 $\boldsymbol{d}_0^{\mathrm{T}}$,由(7.3.4)式及 $\lambda_1\neq0$ 有

$$\boldsymbol{d}_0^{\mathrm{T}}\boldsymbol{Q}\boldsymbol{d}_1=0.$$

这就是说,为使 $\boldsymbol{d}_1$ 直指 $\bar{\boldsymbol{x}}$,$\boldsymbol{d}_1$ 必须与 $\boldsymbol{d}_0$ 是 $\boldsymbol{Q}$ 共轭的.

综上所述,对于问题(7.3.3),当 $n=2$ 时,从任选初始点 $\boldsymbol{x}_0$ 出发,沿任意下降方向 $\boldsymbol{d}_0$ 作一维搜索得 $\boldsymbol{x}_1$,再从点 $\boldsymbol{x}_1$ 出发,沿 $\boldsymbol{d}_0$ 的 $\boldsymbol{Q}$ 共轭方向 $\boldsymbol{d}_1$ 作一维搜索所得的 $\boldsymbol{x}_2$ 必是极小点 $\bar{\boldsymbol{x}}$,即两次迭代就得到了最优解. 一般地,我们有

定理 7.3.3 设 $\boldsymbol{Q}\in\mathbb{R}^{n\times n}$ 为正定矩阵,$\boldsymbol{d}_0,\boldsymbol{d}_1,\cdots,\boldsymbol{d}_{n-1}$ 是 $\mathbb{R}^n$ 中一组 $\boldsymbol{Q}$ 共轭的非零向量. 对于问题(7.3.3),若从任意点 $\boldsymbol{x}_0\in\mathbb{R}^n$ 出发依次沿 $\boldsymbol{d}_0,\boldsymbol{d}_1,\cdots,\boldsymbol{d}_{n-1}$ 进行一维搜索,则至多经过 n 次迭代可得问题(7.3.3)的最优解.

证明 对于问题(7.3.3),设从 $\boldsymbol{x}_0$ 点出发依次沿方向 $\boldsymbol{d}_0,\boldsymbol{d}_1,\cdots,\boldsymbol{d}_{n-1}$ 进行一维搜索产生的迭代点是

$$\boldsymbol{x}_{k+1}=\boldsymbol{x}_k+\lambda_k\boldsymbol{d}_k,k=0,1,\cdots,n-1, \tag{7.3.6}$$

其中 λ_k 使

$$f(\boldsymbol{x}_k+\lambda_k\boldsymbol{d}_k)=\min_{\lambda\geqslant0}f(\boldsymbol{x}_k+\lambda\boldsymbol{d}_k).$$

由 $\nabla f(\boldsymbol{x})=\boldsymbol{Q}\boldsymbol{x}+\boldsymbol{b}$ 和(7.3.6)式知

$$\nabla f(\boldsymbol{x}_{k+1})=\nabla f(\boldsymbol{x}_k)+\lambda_k\boldsymbol{Q}\boldsymbol{d}_k,k=0,1,\cdots,n-1.$$

反复利用上式知

$$\nabla f(\boldsymbol{x}_{k+1})=\nabla f(\boldsymbol{x}_j)+\sum_{i=j}^{k}\lambda_i\boldsymbol{Q}\boldsymbol{d}_i,j=0,1,\cdots,k,$$

从而

$$\nabla f(\boldsymbol{x}_{k+1})^{\mathrm{T}}\boldsymbol{d}_j=\nabla f(\boldsymbol{x}_j)^{\mathrm{T}}\boldsymbol{d}_j+\sum_{i=j}^{k}\lambda_i\boldsymbol{d}_i^{\mathrm{T}}\boldsymbol{Q}\boldsymbol{d}_j,j=0,1,\cdots,k.$$

因为 $\boldsymbol{d}_0,\boldsymbol{d}_1,\cdots,\boldsymbol{d}_{n-1}$ 是 $\boldsymbol{Q}$ 共轭的,所以

$$\begin{aligned}\nabla f(\boldsymbol{x}_{k+1})^{\mathrm{T}}\boldsymbol{d}_j &= \nabla f(\boldsymbol{x}_j)^{\mathrm{T}}\boldsymbol{d}_j + \lambda_j \boldsymbol{d}_j^{\mathrm{T}}\boldsymbol{Q}\boldsymbol{d}_j \\ &= [\nabla f(\boldsymbol{x}_j) + \lambda_j \boldsymbol{Q}\boldsymbol{d}_j]^{\mathrm{T}}\boldsymbol{d}_j,\end{aligned}$$

即知

$$\nabla f(\boldsymbol{x}_{k+1})^{\mathrm{T}}\boldsymbol{d}_j = \nabla f(\boldsymbol{x}_{j+1})^{\mathrm{T}}\boldsymbol{d}_j, j=0,1,\cdots,k. \tag{7.3.7}$$

另外，由于 λ_j 是一维搜索的最优步长，因此

$$\nabla f(\boldsymbol{x}_{j+1})^{\mathrm{T}}\boldsymbol{d}_j = 0, j=0,1,\cdots,n-1.$$

于是，由(7.3.7)式知

$$\nabla f(\boldsymbol{x}_{k+1})^{\mathrm{T}}\boldsymbol{d}_j = 0, k=0,1,\cdots n-1; j=0,1,\cdots,k. \tag{7.3.8}$$

特别地，当 $k=n-1$ 时，有

$$\nabla f(\boldsymbol{x}_n)^{\mathrm{T}}\boldsymbol{d}_j = 0, j=0,1,\cdots,n-1. \tag{7.3.9}$$

因为 $\boldsymbol{d}_0,\boldsymbol{d}_1,\cdots,\boldsymbol{d}_{n-1}$ 是一组 $\boldsymbol{Q}$ 共轭方向，所以由定理 7.3.1 知，它们是线性无关的. 从而由(7.3.9)式及定理 7.3.2 知，$\nabla f(\boldsymbol{x}_n)=\boldsymbol{0}$，即知 $\boldsymbol{x}_n$ 是问题(7.3.3)的最优解.

这说明，至多经过 n 次迭代必定得到(7.3.3)的最优解. □

通常，我们把从点 $\boldsymbol{x}_0 \in \mathbb{R}^n$ 出发、依次沿某组共轭方向进行一维搜索来求解无约束最优化问题(7.1.1)的方法称为共轭方向法(conjugate direction method).

算法 7-4(共轭方向法)

给定一个正定矩阵 $\boldsymbol{Q} \in \mathbb{R}^{n\times n}$.

Step 1　选取初始数据. 选取初始点 $\boldsymbol{x}_0$，给定允许误差 $\varepsilon>0$.

Step 2　选取初始搜索方向. 计算 $\nabla f(\boldsymbol{x}_0)$，求出 $\boldsymbol{d}_0$，使 $\nabla f(\boldsymbol{x}_0)^{\mathrm{T}}\boldsymbol{d}_0<0$，令 $k=0$.

Step 3　检查是否满足终止准则. 若 $\|\nabla f(\boldsymbol{x}_k)\|<\varepsilon$，迭代终止；否则，转 Step 4.

Step 4　进行一维搜索. 求出 λ_k 和 $\boldsymbol{x}_{k+1}$，使得

$$f(\boldsymbol{x}_k+\lambda_k\boldsymbol{d}_k)=\min_{\lambda\geqslant 0} f(\boldsymbol{x}_k+\lambda\boldsymbol{d}_k)$$

$$\boldsymbol{x}_{k+1}=\boldsymbol{x}_k+\lambda_k\boldsymbol{d}_k.$$

Step 5　选取搜索方向. 求 $\boldsymbol{d}_{k+1}$ 使

$$\boldsymbol{d}_{k+1}^{\mathrm{T}}\boldsymbol{Q}\boldsymbol{d}_j=0, j=0,1,\cdots,k.$$

令 $k:=k+1$，返回 Step 3.

如果用共轭方向法求解正定二次函数的无约束最优化问题(7.3.3)(此时算法中的正定矩阵应与二次函数的正定矩阵一致)，那么容易推出迭代公式为

$$\boldsymbol{x}_{k+1}=\boldsymbol{x}_k-\frac{\nabla f(\boldsymbol{x}_k)^{\mathrm{T}}\boldsymbol{d}_k}{\boldsymbol{d}_k^{\mathrm{T}}\boldsymbol{Q}\boldsymbol{d}_k}\boldsymbol{d}_k. \tag{7.3.10}$$

并且由定理 7.3.3 知，经过有限次迭代可以得到最优解，即知共轭方向法具有二次终止性.

7.3.2 共轭梯度法

在共轭方向法中，如果取初始的搜索方向

$$d_0 = -\nabla f(x_0),$$

而以下各共轭方向 d_k 由第 k 次迭代点的负梯度 $-\nabla f(x_k)$ 与已经得到的共轭方向 d_{k-1} 的线性组合来确定，这样就构造了一种具体的共轭方向法. 因为每一个共轭方向都依赖于迭代点处的负梯度，所以称之为共轭梯度法(conjugate gradient method).

先针对问题(7.3.3)，给出共轭梯度法的推导.

给定初始点 $x_0 \in \mathbb{R}^n$，取初始搜索方向

$$d_0 = -\nabla f(x_0),$$

从 x_0 出发，沿 d_0 进行一维搜索，得迭代点 x_1，以下按

$$d_{k+1} = -\nabla f(x_{k+1}) + \alpha_k d_k, k = 0,1,\cdots,n-2$$

来构造搜索方向，α_k 的选取应使所产生的 d_{k+1} 与 d_k 是 Q 共轭的. 因为

$$d_{k+1}^{\mathrm{T}} Q d_k = -\nabla f(x_{k+1})^{\mathrm{T}} Q d_k + \alpha_k d_k^{\mathrm{T}} Q d_k,$$

且要使 d_{k+1} 与 d_k 是 Q 共轭的，应有 $d_{k+1}^{\mathrm{T}} Q d_k = 0$，所以

$$\alpha_k = \frac{\nabla f(x_{k+1})^{\mathrm{T}} Q d_k}{d_k^{\mathrm{T}} Q d_k}, k = 0,1,\cdots,n-2.$$

于是，得到 n 个搜索方向 $d_0, d_1, \cdots, d_{n-1}$ 如下：

$$\begin{cases} d_0 = -\nabla f(x_0), \\ d_{k+1} = -\nabla f(x_{k+1}) + \alpha_k d_k, \\ \alpha_k = \dfrac{\nabla f(x_{k+1})^{\mathrm{T}} Q d_k}{d_k^{\mathrm{T}} Q d_k}. \end{cases} \tag{7.3.11}$$

定理 7.3.4 由(7.3.11)式构造出来的 $d_0, d_1, \cdots, d_{n-1}$ 是一组 Q 共轭方向.

证明 要证 $d_i^{\mathrm{T}} Q d_j = 0, 0 \leqslant i < j \leqslant n-1$. 对 j 用归纳法.

当 $j=1$ 时，由 α_0 的定义知 $d_0^{\mathrm{T}} Q d_1 = 0$. 假设当 $j \leqslant k$ 时，$d_i^{\mathrm{T}} Q d_j = 0 (0 \leqslant i < j \leqslant k)$，下面证明

$$d_i^{\mathrm{T}} Q d_{k+1} = 0, 0 \leqslant i \leqslant k.$$

若 $i = k$，则由(7.3.11)式知

$$d_k^{\mathrm{T}} Q d_{k+1} = -d_k^{\mathrm{T}} Q \nabla f(x_{k+1}) + \frac{\nabla f(x_{k+1})^{\mathrm{T}} Q d_k}{d_k^{\mathrm{T}} Q d_k} d_k^{\mathrm{T}} Q d_k = 0.$$

若 $i < k$，由归纳假设，$d_0, d_1, \cdots, d_k$ 是一组 Q 共轭方向，于是

$$d_i^{\mathrm{T}} Q d_{k+1} = -d_i^{\mathrm{T}} Q \nabla f(x_{k+1}) + \alpha_k d_i^{\mathrm{T}} Q d_k = -d_i^{\mathrm{T}} Q \nabla f(x_{k+1}), \tag{7.3.12}$$

由于 $x_{i+1} = x_i + \lambda_i d_i$，且 $\lambda_i > 0$，因此

$$d_i = \frac{1}{\lambda_i}(x_{i+1} - x_i),$$

代入(7.3.12)式得

$$\begin{aligned} d_i^{\mathrm{T}} Q d_{k+1} &= -\nabla f(x_{k+1})^{\mathrm{T}} Q d_i = -\frac{1}{\lambda_i} \nabla f(x_{k+1})^{\mathrm{T}} Q(x_{i+1} - x_i) \\ &= -\frac{1}{\lambda_i} \nabla f(x_{k+1})^{\mathrm{T}} (\nabla f(x_{i+1}) - \nabla f(x_i)), \end{aligned} \tag{7.3.13}$$

由(7.3.11)式知,$\nabla f(x_i) = -d_i + \alpha_{i-1} d_{i-1}$,从而由(7.3.8)式知,对一切 $i=1$, $2,\cdots,k$,有

$$\nabla f(x_{k+1})^{\mathrm{T}} \nabla f(x_i) = -\nabla f(x_{k+1})^{\mathrm{T}} d_i + \alpha_{i-1} \nabla f(x_{k+1})^{\mathrm{T}} d_{i-1} = 0.$$

而 $i=0$ 时,由(7.3.8)式知

$$\nabla f(x_{k+1})^{\mathrm{T}} \nabla f(x_0) = -\nabla f(x_{k+1})^{\mathrm{T}} d_0 = 0.$$

因此,由(7.3.13)式有 $d_i^{\mathrm{T}} Q d_{k+1} = 0 (i < k)$. □

用共轭梯度法求解问题(7.3.3)时,从点 x_k 出发沿 d_k 进行一维搜索得到的迭代公式为(7.3.10),即知最优步长

$$\lambda_k = -\frac{\nabla f(x_k)^{\mathrm{T}} d_k}{d_k^{\mathrm{T}} Q d_k}.$$

于是,由(7.3.11)中的第2式有

$$\begin{aligned} \lambda_k &= -\frac{\nabla f(x_k)^{\mathrm{T}} (-\nabla f(x_k) + \alpha_{k-1} d_{k-1})}{d_k^{\mathrm{T}} Q d_k} \\ &= \frac{\nabla f(x_k)^{\mathrm{T}} \nabla f(x_k) - \alpha_{k-1} \nabla f(x_k)^{\mathrm{T}} d_{k-1}}{d_k^{\mathrm{T}} Q d_k}, \end{aligned}$$

即

$$\lambda_k = \frac{\nabla f(x_k)^{\mathrm{T}} \nabla f(x_k)}{d_k^{\mathrm{T}} Q d_k}. \tag{7.3.14}$$

这就是用共轭梯度法求解问题(7.3.3)时最优步长的表达式.

为了把共轭梯度法用于求解问题(7.1.1),我们必须设法消失(7.3.11)的第3式中的矩阵 Q,因为

$$\nabla f(x_{k+1}) = \nabla f(x_k) + \lambda_k Q d_k,$$

所以

$$Q d_k = \frac{1}{\lambda_k}(\nabla f(x_{k+1}) - \nabla f(x_k)),$$

代入(7.3.11)中的第3式即可消去 Q,得

$$\alpha_k = \frac{\nabla f(x_{k+1})^{\mathrm{T}} [\nabla f(x_{k+1}) - \nabla f(x_k)]}{d_k^{\mathrm{T}} [\nabla f(x_{k+1}) - \nabla f(x_k)]}.$$

它取不同形式,就得到求解问题(7.1.1)的不同的共轭梯度法.

根据(7.3.8)式，有

$$\nabla f(\boldsymbol{x}_{k+1})^{\mathrm{T}}\boldsymbol{d}_k=0,$$

$$\begin{aligned}\nabla f(\boldsymbol{x}_k)^{\mathrm{T}}\boldsymbol{d}_k&=-\nabla f(\boldsymbol{x}_k)^{\mathrm{T}}\nabla f(\boldsymbol{x}_k)+\alpha_{k-1}\nabla f(\boldsymbol{x}_k)^{\mathrm{T}}\boldsymbol{d}_{k-1}\\&=-\nabla f(\boldsymbol{x}_k)^{\mathrm{T}}\nabla f(\boldsymbol{x}_k),\end{aligned}$$

$$\nabla f(\boldsymbol{x}_{k+1})^{\mathrm{T}}\nabla f(\boldsymbol{x}_k)=-\nabla f(\boldsymbol{x}_{k+1})^{\mathrm{T}}\boldsymbol{d}_k+\alpha_{k-1}\nabla f(\boldsymbol{x}_{k+1})^{\mathrm{T}}\boldsymbol{d}_{k-1}=0,$$

由此可得：

(1) Fletcher－Reeves 公式(简称 FR 公式)

$$\alpha_k=\frac{\|\nabla f(\boldsymbol{x}_{k+1})\|^2}{\|\nabla f(\boldsymbol{x}_k)\|^2};$$

(2) Dixon－Myers 公式(简称 DM 公式)

$$\alpha_k=-\frac{\nabla f(\boldsymbol{x}_{k+1})^{\mathrm{T}}\nabla f(\boldsymbol{x}_{k+1})}{\boldsymbol{d}_k^{\mathrm{T}}\nabla f(\boldsymbol{x}_k)};$$

(3) Polak－Ribiere－Polyak 公式(简称 PRP 公式)

$$\alpha_k=\frac{\nabla f(\boldsymbol{x}_{k+1})^{\mathrm{T}}[\nabla f(\boldsymbol{x}_{k+1})-\nabla f(\boldsymbol{x}_k)]}{\nabla f(\boldsymbol{x}_k)^{\mathrm{T}}\nabla f(\boldsymbol{x}_k)}.$$

根据 α_k 的上述三种形式，可分别给出 FR 共轭梯度法、DM 共轭梯度法和 PRP 共轭梯度法. 对于目标函数是正定二次函数的无约束最优化问题(7.3.3)和最优一维搜索，这些方法是完全等价的. 但是，对于目标函数是非二次函数的无约束最优化问题(7.1.1)，它们所产生的搜索方向是不同的.

需要指出的是，由于$\mathbb{R}^n$ 中共轭方向最多有 n 个，因此在用上述三种方法求解目标函数为非二次函数的无约束最优化问题(7.1.1)时，在 n 步之后构造的搜索方向不再是共轭的，从而降低了收敛速度. 克服的办法是重设初始点，即把经过 n 次迭代得到的 $\boldsymbol{x}_n$ 作为初始点重新迭代.

算法 7－5(FR 共轭梯度法)

Step 1　选取初始数据. 选取初始点 $\boldsymbol{x}_0$，给定允许误差 $\varepsilon>0$.

Step 2　检查是否满足终止准则. 计算$\nabla f(\boldsymbol{x}_0)$，若 $\|\nabla f(\boldsymbol{x}_0)\|<\varepsilon$ 迭代终止，$\boldsymbol{x}_0$ 为(7.1.1)的近似最优解；否则，转 Step 3.

Step 3　构造初始搜索方向. 计算 $\boldsymbol{d}_0=-\nabla f(\boldsymbol{x}_0)$，$k=0$.

Step 4　进行一维搜索. 求出 λ_k 和 $\boldsymbol{x}_{k+1}$，使得

$$f(\boldsymbol{x}_k+\lambda_k\boldsymbol{d}_k)=\min_{\lambda\geqslant 0}f(\boldsymbol{x}_k+\lambda\boldsymbol{d}_k),$$

$$\boldsymbol{x}_{k+1}=\boldsymbol{x}_k+\lambda_k\boldsymbol{d}_k.$$

Step 5　检查是否满足终止准则. 计算$\nabla f(\boldsymbol{x}_{k+1})$，若 $\|\nabla f(\boldsymbol{x}_{k+1})\|<\varepsilon$，迭代终止，$\boldsymbol{x}_{k+1}$为(7.1.1)的近似最优解；否则，转 Step 6.

Step 6　检查迭代次数. 若 $k+1=n$，令 $\boldsymbol{x}_0:=\boldsymbol{x}_n$，返回 Step 3；否则，转 Step 7.

Step 7　构造共轭方向. 用 FR 公式取

$$d_{k+1} = -\nabla f(x_{k+1}) + \alpha_k d_k, \alpha_k = \frac{\| \nabla f(x_{k+1}) \|^2}{\| \nabla f(x_k) \|^2},$$

令 $k := k+1$,返回 Step 4.

如果算法 7-5 的 Step 7 中 α_k 的形式改为 DM 公式或 PRP 公式,则分别得到 DM 共轭梯度法和 PRP 共轭梯度法.

由于共轭梯度法的迭代过程中包含有最速下降法的迭代步骤,因此在较弱的条件下会保证共轭梯度法的收敛性.

定理 7.3.5 设 $f:\mathbb{R}^n \to \mathbb{R}$ 具有一阶连续偏导数,$x_0 \in \mathbb{R}^n$,记 $\alpha = f(x_0)$,并假设水平集 $S(f,\alpha)$ 有界. 若 $\{x_k\}$ 是由共轭梯度法(包括任何一种仅仅与算法 7-5 中 α_k 的形式不同的共轭梯度法)解问题(7.1.1)所产生的点列,则

(1) 当 $\{x_k\}$ 是有穷点列时,其最后一个点是 f 的平稳点;

(2) 当 $\{x_k\}$ 是无穷点列时,它必有极限点,并且其任一极限点都是 f 的平稳点.

证明 (1) 当 $\{x_k\}$ 是有穷点列时,由算法 7-5 的终止条件可知,其最后一个点 $\bar{x}$ 满足 $\nabla f(\bar{x}) = \mathbf{0}$,即 $\bar{x}$ 是 f 的平稳点.

(2) 当 $\{x_k\}$ 是无穷点列时,则

$$\nabla f(x_k) \neq \mathbf{0}, d_k \neq \mathbf{0}, k = 0,1,2,\cdots.$$

因为 x_k 是从点 x_{k-1} 出发沿 d_{k-1} 进行最优一维搜索得到的,所以

$$\nabla f(x_k)^{\mathrm{T}} d_{k-1} = 0.$$

从而

$$\begin{aligned}\nabla f(x_k)^{\mathrm{T}} d_k &= \nabla f(x_k)^{\mathrm{T}}(-\nabla f(x_k) + \alpha_{k-1} d_{k-1}) \\ &= -\nabla f(x_k)^{\mathrm{T}} \nabla f(x_k) < 0,\end{aligned}$$

即 d_k 为 f 在点 x_k 处的下降方向. 而 $x_{k+1} = x_k + \lambda_k d_k$,故

$$f(x_{k+1}) < f(x_k), k = 0,1,2,\cdots.$$

又由于 f 连续,且 $S(f,\alpha)$ 为有界闭集,因此,f 在 $S(f,\alpha)$ 上有界,即 $\{f(x_k)\}$ 为单调减小的有界数列,于是 $\{f(x_k)\}$ 存在极限.

设 $\{x_{k_i}\}$ 是由初始点 x_0 每进行 n 步迭代后按最速下降法得到的点组成的点列,则 $\{x_{k_i}\} \subseteq \{x_k\} \subseteq S(f,\alpha)$,即 $\{x_{k_i}\}$ 为有界点列,从而有收敛的子列,为方便计,可设 $\{x_{k_i}\}$ 收敛,且

$$\lim_{k_i \to \infty} x_{k_i} = \bar{x}.$$

由最速下降法的收敛定理 7.1.1 的证明,可得 $\nabla f(\bar{x}) = \mathbf{0}$,所以 $\bar{x}$ 是 f 的平稳点.

□

可以证明:共轭梯度法产生的点列 $\{x_k\}$ 是 n 步二阶收敛的,即

$$\limsup_{k \to \infty} \frac{\| x_{k+n} - \bar{x} \|}{\| x_k - \bar{x} \|^2} \leqslant q < \infty.$$

详细的证明参见文献[16].

例 7.3.1 用 FR 共轭梯度法求解问题(7.2.6),仍取初始点 $\boldsymbol{x}_0=(0,0)^{\mathrm{T}}$,允许误差 $\varepsilon=0.1$.

解 因为

$$\nabla f(\boldsymbol{x})=(-2(1-x_1)-8(x_2-x_1^2)x_1,4(x_2-x_1^2))^{\mathrm{T}},$$

所以

$$\nabla f(\boldsymbol{x}_0)=(-2,0)^{\mathrm{T}},\ \|\nabla f(\boldsymbol{x}_0)\|=2>\varepsilon.$$

令 $\boldsymbol{d}_0=-\nabla f(\boldsymbol{x}_0)=(2,0)^{\mathrm{T}}$,从点 $\boldsymbol{x}_0$ 出发,沿 $\boldsymbol{d}_0$ 进行一维搜索,得

$$\lambda_0=1/4,\boldsymbol{x}_1=\boldsymbol{x}_0+\lambda_0\boldsymbol{d}_0=(1/2,0)^{\mathrm{T}}.$$

从而

$$\nabla f(\boldsymbol{x}_1)=(0,-1)^{\mathrm{T}},\ \|\nabla f(\boldsymbol{x}_1)\|=1>\varepsilon.$$

由 FR 公式有

$$\alpha_0=\frac{\|\nabla f(\boldsymbol{x}_1)\|^2}{\|\nabla f(\boldsymbol{x}_0)\|^2}=\frac{1}{4},$$

因此,新的搜索方向为

$$\boldsymbol{d}_1=-\nabla f(\boldsymbol{x}_1)+\alpha_0\boldsymbol{d}_0=(1/2,1)^{\mathrm{T}}.$$

从点 $\boldsymbol{x}_1$ 出发沿 $\boldsymbol{d}_1$ 进行一维搜索,得

$$\lambda_1=1,\boldsymbol{x}_2=\boldsymbol{x}_1+\lambda_1\boldsymbol{d}_1=(1,1)^{\mathrm{T}}.$$

此时

$$\nabla f(\boldsymbol{x}_2)=(0,0)^{\mathrm{T}},\ \|\nabla f(\boldsymbol{x}_2)\|=0<\varepsilon.$$

得问题(7.2.6)的最优解为 $\boldsymbol{x}_2=(1,1)^{\mathrm{T}}$. □

7.4 变度量法

变度量法是近 50 年发展起来的一种求解无约束最优化问题的最有效的方法,特别是对于高维问题具有明显的优越性.

7.4.1 拟 Newton 条件

前面介绍的最速下降法和阻尼 Newton 法,它们的迭代公式可以统一为

$$\begin{cases}\boldsymbol{x}_{k+1}=\boldsymbol{x}_k+\lambda_k\boldsymbol{d}_k,\\ \boldsymbol{d}_k=-\boldsymbol{G}_k\nabla f(\boldsymbol{x}_k),\end{cases}\tag{7.4.1}$$

其中 $\boldsymbol{G}_k\in\mathbb{R}^{n\times n}$,$\lambda_k$ 是从点 $\boldsymbol{x}_k$ 出发沿 $\boldsymbol{d}_k$ 进行一维搜索的最优步长. 当 $\boldsymbol{G}_k=\boldsymbol{I}_n$($n$ 阶单位矩阵)时,(7.4.1)式即为最速下降法的迭代公式;当 $\boldsymbol{G}_k=[\nabla^2 f(\boldsymbol{x}_k)]^{-1}$ 时,(7.4.1)式就是阻尼 Newton 法的迭代公式. 因此,如果能够使 $\boldsymbol{G}_k$ 的选取既不

需要计算 Hesse 矩阵及其逆矩阵，又能很好地近似于$[\nabla^2 f(\boldsymbol{x}_k)]^{-1}$，则由(7.4.1)式确定的迭代算法将会保持 Newton 法收敛速度快的优点，同时又具有计算简单的特性.

设问题(7.1.1)中目标函数f具有二阶连续偏导数，且$\nabla^2 f(\boldsymbol{x}_k)$正定. 为使由(7.4.1)中第 2 式确定的搜索方向是f在点$\boldsymbol{x}_k$处的下降方向，根据定理 1.2.3，应当要求

$$\nabla f(\boldsymbol{x}_k)^{\mathrm{T}}\boldsymbol{d}_k < 0,$$

或即

$$\nabla f(\boldsymbol{x}_k)^{\mathrm{T}}\boldsymbol{G}_k \nabla f(\boldsymbol{x}_k) > 0,$$

所以我们应要求(7.4.1)式中的$\boldsymbol{G}_k$是正定矩阵.

设在第$k+1$次迭代后得到$\boldsymbol{x}_{k+1}$，将f在点$\boldsymbol{x}_{k+1}$处作 Taylor 展开，取二阶近似，得

$$f(\boldsymbol{x}) \approx f(\boldsymbol{x}_{k+1}) + \nabla f(\boldsymbol{x}_{k+1})^{\mathrm{T}}(\boldsymbol{x}-\boldsymbol{x}_{k+1}) + \frac{1}{2}(\boldsymbol{x}-\boldsymbol{x}_{k+1})^{\mathrm{T}}\nabla^2 f(\boldsymbol{x}_{k+1})(\boldsymbol{x}-\boldsymbol{x}_{k+1}),$$

对上式两边求梯度，有

$$\nabla f(\boldsymbol{x}) \approx \nabla f(\boldsymbol{x}_{k+1}) + \nabla^2 f(\boldsymbol{x}_{k+1})(\boldsymbol{x}-\boldsymbol{x}_{k+1}),$$

令$\boldsymbol{x}=\boldsymbol{x}_k$，得到

$$\nabla f(\boldsymbol{x}_{k+1}) - \nabla f(\boldsymbol{x}_k) \approx \nabla^2 f(\boldsymbol{x}_{k+1})(\boldsymbol{x}_{k+1}-\boldsymbol{x}_k),$$

即

$$[\nabla^2 f(\boldsymbol{x}_{k+1})]^{-1}[\nabla f(\boldsymbol{x}_{k+1}) - \nabla f(\boldsymbol{x}_k)] \approx \boldsymbol{x}_{k+1}-\boldsymbol{x}_k.$$

易知，当f为正定二次函数时，上式成为等式，即

$$[\nabla^2 f(\boldsymbol{x}_{k+1})]^{-1}[\nabla f(\boldsymbol{x}_{k+1}) - \nabla f(\boldsymbol{x}_k)] = \boldsymbol{x}_{k+1}-\boldsymbol{x}_k. \tag{7.4.2}$$

因为具有正定 Hesse 矩阵的函数在极小点附近可用二次函数很好地近似，所以如果我们迫使$\boldsymbol{G}_{k+1}$满足类似于(7.4.2)式的关系式，即令

$$\boldsymbol{G}_{k+1}[\nabla f(\boldsymbol{x}_{k+1}) - \nabla f(\boldsymbol{x}_k)] = \boldsymbol{x}_{k+1}-\boldsymbol{x}_k, \tag{7.4.3}$$

则$\boldsymbol{G}_{k+1}$就可以很好地近似于$[\nabla f(\boldsymbol{x}_{k+1})]^{-1}$. 因此称(7.4.3)式为拟 Newton 条件(quasi-Newton condition).

为方便起计，记

$$\begin{aligned}
\Delta\boldsymbol{x}_k &= \boldsymbol{x}_{k+1}-\boldsymbol{x}_k,\\
\Delta\boldsymbol{g}_k &= \nabla f(\boldsymbol{x}_{k+1}) - \nabla f(\boldsymbol{x}_k),\\
\Delta\boldsymbol{G}_k &= \boldsymbol{G}_{k+1}-\boldsymbol{G}_k,
\end{aligned}$$

并称$\Delta\boldsymbol{G}_k$为校正矩阵(correction matrix)，则拟 Newton 条件可以写成

$$\Delta\boldsymbol{G}_k\Delta\boldsymbol{g}_k = \Delta\boldsymbol{x}_k - \boldsymbol{G}_k\Delta\boldsymbol{g}_k. \tag{7.4.4}$$

综上所述，我们得到如下的一类算法.

算法 7－6（拟 Newton 法）

Step 1　选取初始数据. 选取初始点 $\boldsymbol{x}_0 \in \mathbb{R}^n$ 和初始矩阵 $\boldsymbol{G}_0$，要求 $\boldsymbol{G}_0$ 为正定矩阵（可取 $\boldsymbol{G}_0 = \boldsymbol{I}_n$），给定允许误差 $\varepsilon > 0$，令 $k = 0$.

Step 2　检查是否满足终止准则. 计算 $\nabla f(\boldsymbol{x}_k)$，若 $\|\nabla f(\boldsymbol{x}_k)\| < \varepsilon$，迭代终止，$\boldsymbol{x}_k$ 为问题(7.1.1)的近似最优解；否则，转 Step 3.

Step 3　构造搜索方向. 令 $\boldsymbol{d}_k = -\boldsymbol{G}_k \nabla f(\boldsymbol{x}_k)$.

Step 4　进行一维搜索. 求 λ_k 和 $\boldsymbol{x}_{k+1}$，使得

$$f(\boldsymbol{x}_k + \lambda_k \boldsymbol{d}_k) = \min_{\lambda \geqslant 0} f(\boldsymbol{x}_k + \lambda \boldsymbol{d}_k),$$

$$\boldsymbol{x}_{k+1} = \boldsymbol{x}_k + \lambda_k \boldsymbol{d}_k.$$

Step 5　产生校正矩阵. 求出满足(7.4.4)的校正矩阵 $\Delta\boldsymbol{G}_k$，令 $\boldsymbol{G}_{k+1} = \boldsymbol{G}_k + \Delta\boldsymbol{G}_k$，$k := k+1$，返回 Step 2.

由于这种算法的搜索方向是根据拟 Newton 条件(7.4.4)构造出来的，因此称之为拟 Newton 法(quasi-Newton method). 因为满足拟 Newton 条件(7.4.4)中的 $\Delta\boldsymbol{G}_k$ 并不是惟一的，所以拟 Newton 法是一族算法.

7.4.2　Q 度量意义下的最速下降方向

设 $f: \mathbb{R}^n \to \mathbb{R}$ 可微，由 1.2 节知，在向量范数为

$$\|\boldsymbol{x}\| = \sqrt{\boldsymbol{x}^{\mathrm{T}}\boldsymbol{x}}$$

的度量意义下，f 在点 $\boldsymbol{x}_k$ 处的最速下降方向为 $-\nabla f(\boldsymbol{x}_k)$. 但是，如果把 $\mathbb{R}^n$ 中的向量范数定义为

$$\|\boldsymbol{x}\|_Q = \sqrt{\boldsymbol{x}^{\mathrm{T}}\boldsymbol{Q}\boldsymbol{x}},$$

其中 $\boldsymbol{Q} \in \mathbb{R}^{n \times n}$ 为正定矩阵，那么，从点 $\boldsymbol{x}_k$ 出发在上述 $\boldsymbol{Q}$ 度量意义下沿哪个方向 $\boldsymbol{d}$ 搜索，f 下降得最快？

设 $\boldsymbol{d} \in \mathbb{R}^n$ 是 $\boldsymbol{Q}$ 度量意义下的单位向量，即

$$\|\boldsymbol{d}\|_Q = \sqrt{\boldsymbol{d}^{\mathrm{T}}\boldsymbol{Q}\boldsymbol{d}} = 1.$$

由定理 1.2.2 知，f 从点 $\boldsymbol{x}_k$ 出发沿 $\boldsymbol{d}$ 的变化率（方向导数）为

$$\frac{\partial f(\boldsymbol{x}_k)}{\partial \boldsymbol{d}} = \nabla f(\boldsymbol{x}_k)^{\mathrm{T}} \frac{\boldsymbol{d}}{\|\boldsymbol{d}\|_Q}.$$

因为 f 从点 $\boldsymbol{x}_k$ 出发下降最快的方向就是沿此方向变化率取最小值的方向，所以，最优化问题

$$\begin{cases} \min \nabla f(\boldsymbol{x}_k)^{\mathrm{T}}\boldsymbol{d}; \\ \text{s.t.} \quad \|\boldsymbol{d}\|_Q = 1 \end{cases} \tag{7.4.5}$$

的解就是从点 $\boldsymbol{x}_k$ 出发使 f 下降最快的方向.

由于 $\boldsymbol{Q}$ 为正定矩阵，因此存在满秩矩阵 $\boldsymbol{L}$，使 $\boldsymbol{Q} = \boldsymbol{L}^{\mathrm{T}}\boldsymbol{L}$，从而

$$d^{\mathrm{T}}Qd = d^{\mathrm{T}}L^{\mathrm{T}}Ld = (Ld)^{\mathrm{T}}(Ld),$$

$$\nabla f(x_k)^{\mathrm{T}}d = \nabla f(x_k)^{\mathrm{T}}L^{-1}Ld = [(L^{-1})^{\mathrm{T}}\nabla f(x_k)]^{\mathrm{T}}(Ld).$$

令 $p = Ld$,则问题(7.4.5)可写成

$$\begin{cases}\min & [(L^{-1})^{\mathrm{T}}\nabla f(x_k)]^{\mathrm{T}}p;\\ \text{s.t.} & \|p\| = 1.\end{cases} \tag{7.4.6}$$

根据1.2中Cauchy不等式可知

$$p = -\frac{(L^{-1})^{\mathrm{T}}\nabla f(x_k)}{\|(L^{-1})^{\mathrm{T}}\nabla f(x_k)\|}$$

是问题(7.4.6)的最小点,从而

$$d = L^{-1}p = -\frac{L^{-1}(L^{-1})^{\mathrm{T}}\nabla f(x_k)}{\|(L^{-1})^{\mathrm{T}}\nabla f(x_k)\|} \tag{7.4.7}$$

是问题(7.4.5)的最小点,即(7.4.7)式所定义的 d 就是在 Q 度量意义下,f 在点 x_k 处下降最快的方向. 于是,方向

$$d_k = -L^{-1}(L^{-1})^{\mathrm{T}}\nabla f(x_k) = -(L^{\mathrm{T}}L)^{-1}\nabla f(x_k),$$

即

$$d_k = -Q^{-1}\nabla f(x_k) \tag{7.4.8}$$

为在 Q 度量意义下,f 在点 x_k 处的最速下降方向.

若令 $Q = \nabla^2 f(x_k)$,则(7.4.8)式为Newton方向,所以阻尼Newton法其实是在 $\nabla^2 f(x_k)$ 度量意义下的最速下降法. 这也说明,由(7.4.1)式确定的迭代算法就是在 G_k^{-1} 度量意义下的最速下降法,只是向量范数的定义(度量)在各次迭代中是变化的,故这类算法又称为变度量法(variable metric method). 由此可知,拟Newton法是一类特殊的变度量法.

7.4.3 DFP法

我们知道,校正矩阵 ΔG_k 的不同取法可以产生不同的拟Newton法. 下面来考虑 ΔG_k 的一种比较简单的形式:

$$\Delta G_k = \Delta x_k p_k^{\mathrm{T}} - G_k \Delta g_k q_k^{\mathrm{T}}, \tag{7.4.9}$$

其中 $p_k, q_k \in \mathbb{R}^n$ 是待定的向量. 于是:

$$\Delta G_k \Delta g_k = (p_k^{\mathrm{T}}\Delta g_k)\Delta x_k - (q_k^{\mathrm{T}}\Delta g_k)G_k\Delta g_k,$$

要使 ΔG_k 满足(7.4.4)式,即要求

$$p_k^{\mathrm{T}}\Delta g_k = 1, q_k^{\mathrm{T}}\Delta g_k = 1. \tag{7.4.10}$$

为了选择满足(7.4.10)式的向量 p_k 和 q_k,考虑到 ΔG_k 应有对称性,可简单地取

$$p_k = \alpha_k \Delta x_k, q_k = \beta_k G_k \Delta g_k. \tag{7.4.11}$$

把上式代入(7.4.10)式得

$$\alpha_k \Delta \boldsymbol{x}_k^{\mathrm{T}} \Delta \boldsymbol{g}_k = 1, \beta_k \Delta \boldsymbol{g}_k^{\mathrm{T}} \boldsymbol{G}_k \Delta \boldsymbol{g}_k = 1,$$

于是,当 $\Delta \boldsymbol{x}_k^{\mathrm{T}} \Delta \boldsymbol{g}_k \neq 0, \Delta \boldsymbol{g}_k^{\mathrm{T}} \boldsymbol{G}_k \Delta \boldsymbol{g}_k \neq 0$ 时,有

$$\alpha_k = \frac{1}{\Delta \boldsymbol{x}_k^{\mathrm{T}} \Delta \boldsymbol{g}_k}, \beta_k = \frac{1}{\Delta \boldsymbol{g}_k^{\mathrm{T}} \boldsymbol{G}_k \Delta \boldsymbol{g}_k},$$

代入(7.4.11)式和(7.4.9)式,得到校正矩阵的表达式

$$\Delta \boldsymbol{G}_k = \frac{\Delta \boldsymbol{x}_k \Delta \boldsymbol{x}_k^{\mathrm{T}}}{\Delta \boldsymbol{x}_k^{\mathrm{T}} \Delta \boldsymbol{g}_k} - \frac{\boldsymbol{G}_k \Delta \boldsymbol{g}_k \Delta \boldsymbol{g}_k^{\mathrm{T}} \boldsymbol{G}_k}{\Delta \boldsymbol{g}_k^{\mathrm{T}} \boldsymbol{G}_k \Delta \boldsymbol{g}_k},$$

进而得到矩阵的迭代公式

$$\boldsymbol{G}_{k+1} = \boldsymbol{G}_k + \frac{\Delta \boldsymbol{x}_k \Delta \boldsymbol{x}_k^{\mathrm{T}}}{\Delta \boldsymbol{x}_k^{\mathrm{T}} \Delta \boldsymbol{g}_k} - \frac{\boldsymbol{G}_k \Delta \boldsymbol{g}_k \Delta \boldsymbol{g}_k^{\mathrm{T}} \boldsymbol{G}_k}{\Delta \boldsymbol{g}_k^{\mathrm{T}} \boldsymbol{G}_k \Delta \boldsymbol{g}_k}. \tag{7.4.12}$$

上述给出的选取 $\Delta \boldsymbol{G}_k$ 的方法是 Davidon(1959)首先提出,后又被 Fletcher 和 Powell(1963)改进简化而成的,故(7.4.12)式叫做 DFP 公式. 把由 DFP 公式产生的矩阵 $\boldsymbol{G}_k$ 代入(7.4.1)式得到的搜索方向叫做 DFP 方向,而每次迭代都用 DFP 方向进行一维搜索的拟 Newton 法就称为 DFP 法.

算法 7-7(DFP 法)

Step 1　选取初始数据. 选取初始点 $\boldsymbol{x}_0$,和初始矩阵 $\boldsymbol{G}_0 = \boldsymbol{I}_n$,给定允许误差 $\varepsilon > 0$.

Step 2　检查是否满足终止准则. 计算 $\nabla f(\boldsymbol{x}_0)$,若 $\|\nabla f(\boldsymbol{x}_0)\| < \varepsilon$,迭代终止,$\boldsymbol{x}_0$ 为问题(7.1.1)的近似最优解;否则,转 Step 3.

Step 3　构造初始 DFP 方向. 取 $\boldsymbol{d}_0 = -\nabla f(\boldsymbol{x}_0)$,令 $k=0$.

Step 4　进行一维搜索. 求出 λ_k 和 $\boldsymbol{x}_{k+1}$,使得

$$\begin{cases} f(\boldsymbol{x}_k + \lambda_k \boldsymbol{d}_k) = \min\limits_{\lambda \geqslant 0} f(\boldsymbol{x}_k + \lambda \boldsymbol{d}_k), \\ \boldsymbol{x}_{k+1} = \boldsymbol{x}_k + \lambda_k \boldsymbol{d}_k. \end{cases}$$

Step 5　检查是否满足终止准则. 计算 $\nabla f(\boldsymbol{x}_{k+1})$,若 $\|\nabla f(\boldsymbol{x}_{k+1})\| < \varepsilon$,迭代终止,$\boldsymbol{x}_{k+1}$ 为问题(7.1.1)的近似最优解;否则,转 Step 6.

Step 6　检查迭代次数. 若 $k+1=n$,令 $\boldsymbol{x}_0 := \boldsymbol{x}_n$,返回 Step 3;否则,转 Step 7.

Step 7　构造 DFP 方向. 用 DFP 公式

$$\boldsymbol{G}_{k+1} = \boldsymbol{G}_k + \frac{\Delta \boldsymbol{x}_k \Delta \boldsymbol{x}_k^{\mathrm{T}}}{\Delta \boldsymbol{x}_k^{\mathrm{T}} \Delta \boldsymbol{g}_k} - \frac{\boldsymbol{G}_k \Delta \boldsymbol{g}_k \Delta \boldsymbol{g}_k^{\mathrm{T}} \boldsymbol{G}_k}{\Delta \boldsymbol{g}_k^{\mathrm{T}} \boldsymbol{G}_k \Delta \boldsymbol{g}_k}$$

算出 $\boldsymbol{G}_{k+1}$,取 $\boldsymbol{d}_{k+1} = -\boldsymbol{G}_{k+1} \nabla f(\boldsymbol{x}_{k+1})$,令 $k := k+1$,返回 Step 4.

在上述算法中,当迭代进行 n 次仍未终止,则要把 $\boldsymbol{x}_n$ 作为初始点重新开始迭代,这是为了减少计算误差的积累而造成的影响.

例 7.4.1　用 DFP 法求解问题(7.2.6),仍取初始点 $\boldsymbol{x}_0 = (0,0)^{\mathrm{T}}$,允许误差 $\varepsilon = 0.1$.

解　由于

$$\nabla f(\boldsymbol{x}) = (-2(1-x_1) - 8(x_2 - x_1^2)x_1, 4(x_2 - x_1^2))^{\mathrm{T}},$$

因此$\nabla f(\boldsymbol{x}_0) = (-2,0)^{\mathrm{T}}$. 从而初始搜索方向为

$$\boldsymbol{d}_0 = -\boldsymbol{G}_0 \nabla f(\boldsymbol{x}_0) = (2,0)^{\mathrm{T}}.$$

由例 7.3.1 知,从 $\boldsymbol{x}_0$ 出发沿 $\boldsymbol{d}_0$ 进行一维搜索,得

$$\boldsymbol{x}_1 = (1/2,0)^{\mathrm{T}}, \nabla f(\boldsymbol{x}_1) = (0,-1)^{\mathrm{T}}, \|\nabla f(\boldsymbol{x}_1)\| = 1 > \varepsilon.$$

由

$$\Delta \boldsymbol{x}_0 = \boldsymbol{x}_1 - \boldsymbol{x}_0 = (1/2,0)^{\mathrm{T}}, \Delta \boldsymbol{g}_0 = \nabla f(\boldsymbol{x}_1) - \nabla f(\boldsymbol{x}_0) = (2,-1)^{\mathrm{T}},$$

可得

$$\Delta \boldsymbol{x}_0^{\mathrm{T}} \Delta \boldsymbol{g}_0 = 1, \Delta \boldsymbol{g}_0^{\mathrm{T}} \boldsymbol{G}_0 \Delta \boldsymbol{g}_0 = 5,$$

代入 DFP 公式(7.4.12)式有

$$\begin{aligned}
\boldsymbol{G}_1 &= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} 1/2 \\ 0 \end{bmatrix}\left(\frac{1}{2},0\right) - \frac{1}{5}\begin{bmatrix} 2 \\ -1 \end{bmatrix}(2,-1) \\
&= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} 1/4 & 0 \\ 0 & 0 \end{bmatrix} - \frac{1}{5}\begin{bmatrix} 4 & -2 \\ -2 & 1 \end{bmatrix} \\
&= \frac{1}{20}\begin{bmatrix} 9 & 8 \\ 8 & 16 \end{bmatrix}.
\end{aligned}$$

由此可构造 DFP 方向

$$\boldsymbol{d}_1 = -\boldsymbol{G}_1 \nabla f(\boldsymbol{x}_1) = -\frac{1}{20}\begin{bmatrix} 9 & 8 \\ 8 & 16 \end{bmatrix}\begin{bmatrix} 0 \\ -1 \end{bmatrix} = \frac{1}{5}\begin{bmatrix} 2 \\ 4 \end{bmatrix}.$$

从点 $\boldsymbol{x}_1$ 出发沿 $\boldsymbol{d}_1$ 进行一维搜索,得

$$\lambda_1 = \frac{5}{4}, \boldsymbol{x}_2 = \boldsymbol{x}_1 + \lambda_1 \boldsymbol{d}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

此时,$\nabla f(\boldsymbol{x}_2) = (0,0)^{\mathrm{T}}$, $\|\nabla f(\boldsymbol{x}_2)\| = 0 < \varepsilon$,于是,得问题(7.2.6)的最优解为 $\boldsymbol{x}_2 = (1,1)^{\mathrm{T}}$. □

下面我们证明 DFP 法产生的$\{\boldsymbol{G}_k\}$是正定矩阵序列.

定理 7.4.1 设 $\boldsymbol{G}_k$ 是正定矩阵,且$\nabla f(\boldsymbol{x}_k) \neq \boldsymbol{0}$,则由 DFP 公式(7.4.12)式产生的 $\boldsymbol{G}_{k+1}$ 也是正定矩阵.

证明 先证明 DFP 公式(7.4.12)式中第 2 项和第 3 项的分母为正数.

因为 $\Delta \boldsymbol{x}_k = \boldsymbol{x}_{k+1} - \boldsymbol{x}_k = \lambda_k \boldsymbol{d}_k$,且 λ_k 是从点 $\boldsymbol{x}_k$ 出发沿 $\boldsymbol{d}_k$ 进行一维搜索的最优步长,所以

$$\nabla f(\boldsymbol{x}_{k+1})^{\mathrm{T}} \boldsymbol{d}_k = 0.$$

从而,由(7.4.1)中第 2 式有

$$\begin{aligned}
\Delta \boldsymbol{x}_k^{\mathrm{T}} \Delta \boldsymbol{g}_k &= \lambda_k \boldsymbol{d}_k^{\mathrm{T}} [\nabla f(\boldsymbol{x}_{k+1}) - \nabla f(\boldsymbol{x}_k)] \\
&= -\lambda_k \boldsymbol{d}_k^{\mathrm{T}} \nabla f(\boldsymbol{x}_k) = \lambda_k \nabla f(\boldsymbol{x}_k)^{\mathrm{T}} \boldsymbol{G}_k \nabla f(\boldsymbol{x}_k).
\end{aligned}$$

因$\nabla f(\boldsymbol{x}_k) \neq \boldsymbol{0}$,$\boldsymbol{G}_k$ 正定,故

$$\nabla f(\boldsymbol{x}_k)^{\mathrm{T}} \boldsymbol{G}_k \nabla f(\boldsymbol{x}_k) > 0,$$

而 $\lambda_k > 0$,因此 $\Delta \boldsymbol{x}_k^{\mathrm{T}} \Delta \boldsymbol{g}_k > 0$,即(7.4.12)式中第 2 项的分母为正数. 又由 $\Delta \boldsymbol{x}_k^{\mathrm{T}} \Delta \boldsymbol{g}_k > 0$ 知,$\Delta \boldsymbol{g}_k \neq \boldsymbol{0}$,故由 $\boldsymbol{G}_k$ 正定可知,$\Delta \boldsymbol{g}_k^{\mathrm{T}} \boldsymbol{G}_k \Delta \boldsymbol{g}_k > 0$,即(7.4.12)式中第 3 项的分母为正数.

再证 $\boldsymbol{G}_{k+1}$ 是正定的. 由于 $\boldsymbol{G}_k$ 是对称矩阵,因此由(7.4.12)式知,$\boldsymbol{G}_{k+1}$ 是对称矩阵. 另一方面,对任意非零向量 $\boldsymbol{y} \in \mathbb{R}^n$,有

$$\begin{aligned}\boldsymbol{y}^{\mathrm{T}} \boldsymbol{G}_{k+1} \boldsymbol{y} &= \boldsymbol{y}^{\mathrm{T}} \left[\boldsymbol{G}_k + \frac{\Delta \boldsymbol{x}_k \Delta \boldsymbol{x}_k^{\mathrm{T}}}{\Delta \boldsymbol{x}_k^{\mathrm{T}} \Delta \boldsymbol{g}_k} - \frac{\boldsymbol{G}_k \Delta \boldsymbol{g}_k \Delta \boldsymbol{g}_k^{\mathrm{T}} \boldsymbol{G}_k}{\Delta \boldsymbol{g}_k^{\mathrm{T}} \boldsymbol{G}_k \Delta \boldsymbol{g}_k} \right] \boldsymbol{y} \\ &= \boldsymbol{y}^{\mathrm{T}} \boldsymbol{G}_k \boldsymbol{y} + \frac{(\boldsymbol{y}^{\mathrm{T}} \Delta \boldsymbol{x}_k)^2}{\Delta \boldsymbol{x}_k^{\mathrm{T}} \Delta \boldsymbol{g}_k} - \frac{(\boldsymbol{y}^{\mathrm{T}} \boldsymbol{G}_k \Delta \boldsymbol{g}_k)^2}{\Delta \boldsymbol{g}_k^{\mathrm{T}} \boldsymbol{G}_k \Delta \boldsymbol{g}_k},\end{aligned}$$

因为 $\boldsymbol{G}_k$ 正定,所以存在满秩矩阵 $\boldsymbol{L}_k$,使 $\boldsymbol{G}_k = \boldsymbol{L}_k^{\mathrm{T}} \boldsymbol{L}_k$,令

$$\boldsymbol{p} = \boldsymbol{L}_k \boldsymbol{y}, \boldsymbol{q} = \boldsymbol{L}_k \Delta \boldsymbol{g}_k, \tag{7.4.13}$$

则有

$$\begin{aligned}&\boldsymbol{y}^{\mathrm{T}} \boldsymbol{G}_k \boldsymbol{y} = \boldsymbol{y}^{\mathrm{T}} \boldsymbol{L}_k^{\mathrm{T}} \boldsymbol{L}_k \boldsymbol{y} = (\boldsymbol{L}_k \boldsymbol{y})^{\mathrm{T}} (\boldsymbol{L}_k \boldsymbol{y}) = \| \boldsymbol{p} \|^2, \\ &\Delta \boldsymbol{g}_k^{\mathrm{T}} \boldsymbol{G}_k \Delta \boldsymbol{g}_k = \Delta \boldsymbol{g}_k^{\mathrm{T}} \boldsymbol{L}_k^{\mathrm{T}} \boldsymbol{L}_k \Delta \boldsymbol{g}_k = (\boldsymbol{L}_k \Delta \boldsymbol{g}_k)^{\mathrm{T}} (\boldsymbol{L}_k \Delta \boldsymbol{g}_k) = \| \boldsymbol{q} \|^2, \\ &\boldsymbol{y}^{\mathrm{T}} \boldsymbol{G}_k \Delta \boldsymbol{g}_k = \boldsymbol{y}^{\mathrm{T}} \boldsymbol{L}_k^{\mathrm{T}} \boldsymbol{L}_k \Delta \boldsymbol{g}_k = (\boldsymbol{L}_k \boldsymbol{y})^{\mathrm{T}} (\boldsymbol{L}_k \Delta \boldsymbol{g}_k) = \boldsymbol{p}^{\mathrm{T}} \boldsymbol{q},\end{aligned}$$

于是

$$\boldsymbol{y}^{\mathrm{T}} \boldsymbol{G}_{k+1} \boldsymbol{y} = \| \boldsymbol{p} \|^2 + \frac{(\boldsymbol{y}^{\mathrm{T}} \Delta \boldsymbol{x}_k)^2}{\Delta \boldsymbol{x}_k^{\mathrm{T}} \Delta \boldsymbol{g}_k} - \frac{(\boldsymbol{p}^{\mathrm{T}} \boldsymbol{q})^2}{\| \boldsymbol{q} \|^2},$$

或即

$$\boldsymbol{y}^{\mathrm{T}} \boldsymbol{G}_{k+1} \boldsymbol{y} = \frac{\| \boldsymbol{p} \|^2 \| \boldsymbol{q} \|^2 - (\boldsymbol{p}^{\mathrm{T}} \boldsymbol{q})^2}{\| \boldsymbol{q} \|^2} + \frac{(\boldsymbol{y}^{\mathrm{T}} \Delta \boldsymbol{x}_k)^2}{\Delta \boldsymbol{x}_k^{\mathrm{T}} \Delta \boldsymbol{g}_k}. \tag{7.4.14}$$

以下分两种情况讨论:

(1) $\boldsymbol{p} = \beta \boldsymbol{q} (\beta \neq 0)$,这时(7.4.14)式中第 1 项为 0,并且由(7.4.13)式可知,$\boldsymbol{y} = \beta \Delta \boldsymbol{g}_k$. 将此式代入(7.4.14)式得

$$\boldsymbol{y}^{\mathrm{T}} \boldsymbol{G}_{k+1} \boldsymbol{y} = \frac{(\boldsymbol{y}^{\mathrm{T}} \Delta \boldsymbol{x}_k)^2}{\Delta \boldsymbol{x}_k^{\mathrm{T}} \Delta \boldsymbol{g}_k} = \beta^2 \Delta \boldsymbol{g}_k^{\mathrm{T}} \Delta \boldsymbol{x}_k > 0.$$

(2) $\boldsymbol{p} \neq \beta \boldsymbol{q} (\beta \neq 0)$,此时由 Cauchy 不等式知(7.4.14)式中第 1 项大于 0,而(7.4.14)式中第 2 项显然是非负的,所以也有

$$\boldsymbol{y}^{\mathrm{T}} \boldsymbol{G}_{k+1} \boldsymbol{y} > 0,$$

这就证明了 $\boldsymbol{G}_{k+1}$ 是正定矩阵. □

因为初始矩阵 $\boldsymbol{G}_0 = \boldsymbol{I}_n$ 是正定的,所以由该定理知,DFP 法产生的所有矩阵 $\boldsymbol{G}_k$ 均是正定的.

定理 7.4.2 用 DFP 法求解正定二次函数的无约束最优化问题(7.1.2),若

迭代点是互异的，并且产生的搜索方向依次是 $\boldsymbol{d}_0,\boldsymbol{d}_1,\cdots,\boldsymbol{d}_k(k\leqslant n-1)$，则有

$$\boldsymbol{G}_k\boldsymbol{Q}\boldsymbol{d}_i=\boldsymbol{d}_i,i=0,1,\cdots,k-1, \tag{7.4.15}$$

和

$$\boldsymbol{d}_i^{\mathrm{T}}\boldsymbol{Q}\boldsymbol{d}_j=0,i,j=0,1,\cdots,k,i<j. \tag{7.4.16}$$

证明 首先由迭代点互异的假设知，搜索步长 $\lambda_i>0(i=0,1,\cdots,k)$. 对 k 用数学归纳法证明.

当 $k=1$ 时，由于

$$\Delta\boldsymbol{x}_0=\boldsymbol{x}_1-\boldsymbol{x}_0=\lambda_0\boldsymbol{d}_0,$$
$$\Delta\boldsymbol{g}_0=\nabla f(\boldsymbol{x}_1)-\nabla f(\boldsymbol{x}_0)=(\boldsymbol{Q}\boldsymbol{x}_1+\boldsymbol{b})-(\boldsymbol{Q}\boldsymbol{x}_0+\boldsymbol{b})=\boldsymbol{Q}\Delta\boldsymbol{x}_0,$$

因此

$$\boldsymbol{Q}\boldsymbol{d}_0=\boldsymbol{Q}\left(\frac{1}{\lambda_0}\Delta\boldsymbol{x}_0\right)=\frac{1}{\lambda_0}\Delta\boldsymbol{g}_0,$$

从而，由 DFP 公式(7.4.12)式知

$$\begin{aligned}\boldsymbol{G}_1\boldsymbol{Q}\boldsymbol{d}_0&=\left(\boldsymbol{G}_0+\frac{\Delta\boldsymbol{x}_0\Delta\boldsymbol{x}_0^{\mathrm{T}}}{\Delta\boldsymbol{x}_0^{\mathrm{T}}\Delta\boldsymbol{g}_0}-\frac{\boldsymbol{G}_0\Delta\boldsymbol{g}_0\Delta\boldsymbol{g}_0^{\mathrm{T}}\boldsymbol{G}_0}{\Delta\boldsymbol{g}_0^{\mathrm{T}}\boldsymbol{G}_0\Delta\boldsymbol{g}_0}\right)\frac{1}{\lambda_0}\Delta\boldsymbol{g}_0\\&=\frac{1}{\lambda_0}(\boldsymbol{G}_0\Delta\boldsymbol{g}_0+\Delta\boldsymbol{x}_0-\boldsymbol{G}_0\Delta\boldsymbol{g}_0)=\frac{1}{\lambda_0}\Delta\boldsymbol{x}_0=\boldsymbol{d}_0,\end{aligned}$$

并由此可得

$$\begin{aligned}\boldsymbol{d}_0^{\mathrm{T}}\boldsymbol{Q}\boldsymbol{d}_1&=\boldsymbol{d}_0^{\mathrm{T}}\boldsymbol{Q}(-\boldsymbol{G}_1\nabla f(\boldsymbol{x}_1))=-\nabla f(\boldsymbol{x}_1)^{\mathrm{T}}\boldsymbol{G}_1\boldsymbol{Q}\boldsymbol{d}_0\\&=-\nabla f(\boldsymbol{x}_1)^{\mathrm{T}}\boldsymbol{d}_0=0,\end{aligned}$$

所以，$k=1$ 时，(7.4.15)式和(7.4.16)式成立.

假设对于 $k=m\geqslant1$，(7.4.15)式和(7.4.16)式成立. 下面证明当 $k=m+1$ 时，(7.4.15)式和(7.4.16)式也成立.

由 DFP 公式(7.4.12)有

$$\boldsymbol{G}_{m+1}\boldsymbol{Q}\boldsymbol{d}_i=\boldsymbol{G}_m\boldsymbol{Q}\boldsymbol{d}_i+\frac{\Delta\boldsymbol{x}_m\Delta\boldsymbol{x}_m^{\mathrm{T}}\boldsymbol{Q}\boldsymbol{d}_i}{\Delta\boldsymbol{x}_m^{\mathrm{T}}\Delta\boldsymbol{g}_m}-\frac{\boldsymbol{G}_m\Delta\boldsymbol{g}_m\Delta\boldsymbol{g}_m^{\mathrm{T}}\boldsymbol{G}_m\boldsymbol{Q}\boldsymbol{d}_i}{\Delta\boldsymbol{g}_m^{\mathrm{T}}\boldsymbol{G}_m\Delta\boldsymbol{g}_m}, \tag{7.4.17}$$

并且又有

$$\Delta\boldsymbol{x}_m=\boldsymbol{x}_{m+1}-\boldsymbol{x}_m=\lambda_m\boldsymbol{d}_m, \tag{7.4.18}$$
$$\Delta\boldsymbol{g}_m=\nabla f(\boldsymbol{x}_{m+1})-\nabla f(\boldsymbol{x}_m)=\boldsymbol{Q}\Delta\boldsymbol{x}_m=\lambda_m\boldsymbol{Q}\boldsymbol{d}_m. \tag{7.4.19}$$

当 $i<m$ 时，由上面两式及归纳假设知，对于 $i=0,1,\cdots,m-1$，有

$$\boldsymbol{G}_m\boldsymbol{Q}\boldsymbol{d}_i=\boldsymbol{d}_i,$$
$$\Delta\boldsymbol{x}_m^{\mathrm{T}}\boldsymbol{Q}\boldsymbol{d}_i=\lambda_m\boldsymbol{d}_m^{\mathrm{T}}\boldsymbol{Q}\boldsymbol{d}_i=0,$$
$$\Delta\boldsymbol{g}_m^{\mathrm{T}}\boldsymbol{G}_m\boldsymbol{Q}\boldsymbol{d}_i=\Delta\boldsymbol{g}_m^{\mathrm{T}}\boldsymbol{d}_i=\lambda_m\boldsymbol{d}_m^{\mathrm{T}}\boldsymbol{Q}\boldsymbol{d}_i=0,$$

从而由(7.4.17)式得

$$\boldsymbol{G}_{m+1}\boldsymbol{Q}\boldsymbol{d}_i=\boldsymbol{d}_i,i=0,1,\cdots,m-1. \tag{7.4.20}$$

当 $i=m$ 时,由(7.4.19)式和(7.4.18)式有

$$G_m Q d_m = G_m\left(\frac{1}{\lambda_m}\Delta g_m\right) = \frac{1}{\lambda_m}G_m \Delta g_m,$$

$$\frac{\Delta x_m \Delta x_m^{\mathrm{T}} Q d_m}{\Delta x_m^{\mathrm{T}} \Delta g_m} = \frac{\Delta x_m \Delta x_m^{\mathrm{T}}\left(\frac{1}{\lambda_m}\Delta g_m\right)}{\Delta x_m^{\mathrm{T}} \Delta g_m} = \frac{1}{\lambda_m}\Delta x_m = d_m,$$

$$\frac{G_m \Delta g_m \Delta g_m^{\mathrm{T}} G_m Q d_m}{\Delta g_m^{\mathrm{T}} G_m \Delta g_m} = \frac{G_m \Delta g_m \Delta g_m^{\mathrm{T}} G_m\left(\frac{1}{\lambda_m}\Delta g_m\right)}{\Delta g_m^{\mathrm{T}} G_m \Delta g_m} = \frac{1}{\lambda_m}G_m \Delta g_m.$$

把它们代入(7.4.17)式得到

$$G_{m+1} Q d_m = \frac{1}{\lambda_m}G_m \Delta g_m + d_m - \frac{1}{\lambda_m}G_m \Delta g_m = d_m, \tag{7.4.21}$$

从而由(7.4.20)式和(7.4.21)式知,当 $k=m+1$ 时,(7.4.15)式成立,即

$$G_{m+1} Q d_i = d_i, i=0,1,\cdots,m. \tag{7.4.22}$$

又由归纳假设有

$$d_i^{\mathrm{T}} Q d_j = 0, 0 \leqslant i < j \leqslant m, \tag{7.4.23}$$

并且由 $d_{m+1} = -G_{m+1}\nabla f(x_{m+1})$,有

$$d_i^{\mathrm{T}} Q d_{m+1} = -d_i^{\mathrm{T}} Q G_{m+1}\nabla f(x_{m+1}) = -\nabla f(x_{m+1})^{\mathrm{T}} G_{m+1} Q d_i,$$

从而由(7.4.22)式得

$$d_i^{\mathrm{T}} Q d_{m+1} = -\nabla f(x_{m+1})^{\mathrm{T}} d_i, i=0,1,\cdots,m. \tag{7.4.24}$$

又由(7.4.23)式知,$d_0, d_1, \cdots, d_m$ 是一组 Q 共轭方向. 因此由(7.3.8)式知 $\nabla f(x_{m+1})^{\mathrm{T}} d_i = 0 (i=0,1,\cdots,m)$,于是,由(7.4.24)式得到

$$d_i^{\mathrm{T}} Q d_{m+1} = 0, i=0,1,\cdots,m.$$

根据上式以及(7.4.23)式得知,当 $k=m+1$ 时,(7.4.16)式成立. □

定理 7.4.3 设用 DFP 法求解正定二次函数的无约束最优化问题(7.1.2). 若迭代点是互异的,则至多经过 n 次迭代可得问题(7.1.2)的最优解. 如果总共要进行 n 次迭代算法才终止,则 $G_n = Q^{-1}$.

证明 由定理 7.4.2 中的(7.4.16)式可知,用 DFP 法求解问题(7.1.2)产生的搜索方向 $d_0, d_1, \cdots, d_{n-1}$ 是一组 Q 共轭方向,所以 DFP 法也是一种共轭方向法,从而由定理 7.3.3 知,至多经过 n 次迭代可达(7.1.2)的最优解.

如果要进行 n 次迭代算法才终止,则由(7.4.15)式有

$$G_n Q d_i = d_i, i=0,1,\cdots,n-1, \tag{7.4.25}$$

且由(7.4.16)式知,$d_0, d_1, \cdots, d_{n-1}$ 是一组 Q 共轭方向,因此由定理 7.3.1 知它们是线性无关. 取矩阵

$$D = (d_0, d_1, \cdots, d_{n-1}),$$

则 D 是满秩矩阵,据此,由(7.4.25)式有 $G_n Q D = D$,从而 $G_n Q = I_n$,即 $G_n =$

Q^{-1}. □

上面的定理证明 DFP 法是共轭方向法，因而具有二次终止性. 如果用 DFP 法求解正定二次函数的无约束最优解问题(7.1.2)，算法经过 n 次迭代才终止，则最后的迭代矩阵 $\boldsymbol{G}_n$ 即为目标函数 f 的 Hesse 矩阵的逆矩阵，即知 DFP 法的最后一个搜索方向就是直指问题最优解的 Newton 方向.

定理 7.4.4 设 $f:\mathbb{R}^n\to\mathbb{R}$ 具有一阶连续偏导数，$\boldsymbol{x}_0\in\mathbb{R}^n$，记 $\alpha=f(\boldsymbol{x}_0)$，并假设水平集 $S(f,\alpha)$ 有界. 若 $\{\boldsymbol{x}_k\}$ 是由 DFP 法求解问题(7.1.1)产生的点列，则

(1) 当 $\{\boldsymbol{x}_k\}$ 是有穷点列时，其最后一个点是 f 的平稳点；

(2) 当 $\{\boldsymbol{x}_k\}$ 是无穷点列时，它必有极限点，并且其任一极限点都是 f 的平稳点.

证明 (1) 当 $\{\boldsymbol{x}_k\}$ 是有穷点列时，由 DFP 法的终止条件知，其最后一个点 $\bar{\boldsymbol{x}}$ 满足 $\nabla f(\bar{\boldsymbol{x}})=\boldsymbol{0}$，即 $\bar{\boldsymbol{x}}$ 是 f 的平稳点.

(2) 当 $\{\boldsymbol{x}_k\}$ 是无穷点列时，$\nabla f(\boldsymbol{x}_k)\neq\boldsymbol{0}(k=0,1,2,\cdots)$，又由定理 7.4.1 知，DFP 法产生的所有矩阵 $\boldsymbol{G}_k$ 是正定的$(k=0,1,2,\cdots)$，故

$$\nabla f(\boldsymbol{x}_k)^{\mathrm{T}}\boldsymbol{G}_k\nabla f(\boldsymbol{x}_k)>0.$$

于是，由 $\boldsymbol{d}_k=-\boldsymbol{G}_k\nabla f(\boldsymbol{x}_k)$ 有

$$\nabla f(\boldsymbol{x}_k)^{\mathrm{T}}\boldsymbol{d}_k=-\nabla f(\boldsymbol{x}_k)^{\mathrm{T}}\boldsymbol{G}_k\nabla f(\boldsymbol{x}_k)<0,$$

即 $\boldsymbol{d}_k$ 为 f 在点 $\boldsymbol{x}_k$ 处的下降方向. 从而

$$f(\boldsymbol{x}_{k+1})=f(\boldsymbol{x}_k+\lambda_k\boldsymbol{d}_k)<f(\boldsymbol{x}_k),k=0,1,2,\cdots.$$

以下与定理 7.3.5 的证明类似，可证：$\{\boldsymbol{x}_k\}$ 存在极限点，其任一极限点都是 f 的平稳点. □

还可证明 DFP 法具有超线性收敛性，详细证明参见文献[16,50].

7.4.4 BFGS 法

如果在 7.4.1 节中推导拟 Newton 条件时，一开始不考虑构造逼近 $[\nabla^2 f(\boldsymbol{x}_k)]^{-1}$ 的迭代矩阵 $\boldsymbol{G}_k$，而是考虑逼近 $\nabla^2 f(\boldsymbol{x}_k)$ 的迭代矩阵 $\boldsymbol{B}_k$，则相应于(7.4.1)中第 2 式有

$$\boldsymbol{d}_k=-\boldsymbol{B}_k^{-1}\nabla f(\boldsymbol{x}_k), \tag{7.4.26}$$

并且相应于拟 Newton 条件(7.4.3)有

$$\nabla f(\boldsymbol{x}_{k+1})-\nabla f(\boldsymbol{x}_k)=\boldsymbol{B}_{k+1}(\boldsymbol{x}_{k+1}-\boldsymbol{x}_k),$$

从而，相应于(7.4.4)式有

$$\Delta\boldsymbol{B}_k\Delta\boldsymbol{x}_k=\Delta\boldsymbol{g}_k-\boldsymbol{B}_k\Delta\boldsymbol{x}_k, \tag{7.4.27}$$

其中 $\Delta\boldsymbol{B}_k=\boldsymbol{B}_{k+1}-\boldsymbol{B}_k$. 由此出发，同样可以构造出拟 Newton 法，这类算法中每次迭代的搜索方向 $\boldsymbol{d}_k$ 其实是在 $\boldsymbol{B}_k$ 度量意义下的最速下降方向，故这类算法也是变度量法.

与 7.4.3 节类似,考虑 $\Delta \boldsymbol{B}_k$ 的一种简单形式

$$\Delta \boldsymbol{B}_k = \Delta \boldsymbol{g}_k \boldsymbol{p}_k^{\mathrm{T}} - \boldsymbol{B}_k \Delta \boldsymbol{x}_k \boldsymbol{q}_k^{\mathrm{T}},$$

其中 $\boldsymbol{p}_k, \boldsymbol{q}_k \in \mathbb{R}^n$ 是待定的向量. 同理可得满足(7.4.27)式的校正矩阵

$$\Delta \boldsymbol{B}_k = \frac{\Delta \boldsymbol{g}_k \Delta \boldsymbol{g}_k^{\mathrm{T}}}{\Delta \boldsymbol{g}_k^{\mathrm{T}} \Delta \boldsymbol{x}_k} - \frac{\boldsymbol{B}_k \Delta \boldsymbol{x}_k \Delta \boldsymbol{x}_k^{\mathrm{T}} \boldsymbol{B}_k}{\Delta \boldsymbol{x}_k^{\mathrm{T}} \boldsymbol{B}_k \Delta \boldsymbol{x}_k}.$$

从而得到

$$\boldsymbol{B}_{k+1} = \boldsymbol{B}_k + \frac{\Delta \boldsymbol{g}_k \Delta \boldsymbol{g}_k^{\mathrm{T}}}{\Delta \boldsymbol{g}_k^{\mathrm{T}} \Delta \boldsymbol{x}_k} - \frac{\boldsymbol{B}_k \Delta \boldsymbol{x}_k \Delta \boldsymbol{x}_k^{\mathrm{T}} \boldsymbol{B}_k}{\Delta \boldsymbol{x}_k^{\mathrm{T}} \boldsymbol{B}_k \Delta \boldsymbol{x}_k}. \tag{7.4.28}$$

为了按(7.4.26)式构造出搜索方向 $\boldsymbol{d}_k$,必须求出 $\boldsymbol{B}_k^{-1}$,于是,两次利用 Sherman - Morrison 公式:若 $\boldsymbol{A} \in \mathbb{R}^{n \times n}$ 为满秩矩阵,$\boldsymbol{u}, \boldsymbol{v} \in \mathbb{R}^n$,且 $\boldsymbol{v}^{\mathrm{T}} \boldsymbol{A}^{-1} \boldsymbol{u} \neq -1$,则

$$(\boldsymbol{A} + \boldsymbol{u}\boldsymbol{v}^{\mathrm{T}})^{-1} = \boldsymbol{A}^{-1} - \frac{\boldsymbol{A}^{-1} \boldsymbol{u} \boldsymbol{v}^{\mathrm{T}} \boldsymbol{A}^{-1}}{1 + \boldsymbol{v}^{\mathrm{T}} \boldsymbol{A}^{-1} \boldsymbol{u}}.$$

不难由(7.4.28)式推出

$$\boldsymbol{B}_{k+1}^{-1} = \left(\boldsymbol{I}_n - \frac{\Delta \boldsymbol{x}_k \Delta \boldsymbol{g}_k^{\mathrm{T}}}{\Delta \boldsymbol{x}_k^{\mathrm{T}} \Delta \boldsymbol{g}_k} \right) \boldsymbol{B}_k^{-1} \left(\boldsymbol{I}_n - \frac{\Delta \boldsymbol{x}_k \Delta \boldsymbol{g}_k^{\mathrm{T}}}{\Delta \boldsymbol{x}_k^{\mathrm{T}} \Delta \boldsymbol{g}_k} \right)^{\mathrm{T}} + \frac{\Delta \boldsymbol{x}_k \Delta \boldsymbol{x}_k^{\mathrm{T}}}{\Delta \boldsymbol{x}_k^{\mathrm{T}} \Delta \boldsymbol{g}_k}.$$

令 $\boldsymbol{G}_k = \boldsymbol{B}_k^{-1}$,则由上式有

$$\boldsymbol{G}_{k+1} = \boldsymbol{G}_k + \frac{\Delta \boldsymbol{x}_k \Delta \boldsymbol{x}_k^{\mathrm{T}}}{\Delta \boldsymbol{x}_k^{\mathrm{T}} \Delta \boldsymbol{g}_k} \left(1 + \frac{\Delta \boldsymbol{g}_k^{\mathrm{T}} \boldsymbol{G}_k \Delta \boldsymbol{g}_k}{\Delta \boldsymbol{x}_k^{\mathrm{T}} \Delta \boldsymbol{g}_k} \right) - \frac{1}{\Delta \boldsymbol{x}_k^{\mathrm{T}} \Delta \boldsymbol{g}_k} (\Delta \boldsymbol{x}_k \Delta \boldsymbol{g}_k^{\mathrm{T}} \boldsymbol{G}_k + \boldsymbol{G}_k \Delta \boldsymbol{g}_k \Delta \boldsymbol{x}_k^{\mathrm{T}}). \tag{7.4.29}$$

公式(7.4.29)是 Brayden, Fletcher(1970), Goldfarb(1969)和 Shanno(1970)共同研究的结果,故叫做 BFGS 公式. 把 BFGS 公式产生的矩阵 $\boldsymbol{G}_k$ 构造的搜索方向

$$\boldsymbol{d}_k = -\boldsymbol{G}_k \nabla f(\boldsymbol{x}_k)$$

称为 BFGS 方向. 每一次迭代都用 BFGS 方向进行一维搜索的拟 Newton 法称作 BFGS 法. 于是,只要把 DFP 法的 Step 7 用 DFP 公式构造 DFP 方向改为用 BFGS 公式构造 BFGS 方向,就得到了 BFGS 法.

BFGS 法具有与 DFP 法完全相同的性质,例如,BFGS 法也是共轭方向法,因而具有二次终止性,并且 BFGS 法还具有超线性收敛性.

对于 DFP 法,由于一维搜索的不精确和计算误差的积累可能导致某一次迭代中 $\boldsymbol{G}_k$ 的奇异,而 BFGS 法对一维搜索的精度要求不高,并且由它产生的 $\boldsymbol{G}_k$ 不易变为奇异矩阵. 因此,BFGS 法比 DFP 法具有这一好的数值稳定性,使它更具实用性.

例 7.4.2 用 BFGS 法求解问题(7.2.6),仍取初始点 $\boldsymbol{x}_0 = (0,0)^{\mathrm{T}}$,允许误差 $\varepsilon = 0.1$.

解 由例 7.4.1 知,初始搜索方向

$$\boldsymbol{d}_0 = -\boldsymbol{G}_0 \nabla f(\boldsymbol{x}_0) = (2,0)^{\mathrm{T}},$$

从 x_0 出发沿 d_0 进行一维搜索,得 $x_1=(1/2,0)^T$;并且

$$\nabla f(x_1)=(0,-1)^T,\Delta x_0=(1/2,0)^T,\Delta g_0=(2,-1)^T,$$

从而

$$\Delta x_0^T\Delta g_0=1,\Delta g_0^T G_0\Delta g_0=5,$$

$$\Delta x_0\Delta x_0^T=\begin{bmatrix}1/2\\0\end{bmatrix}\left(\frac{1}{2},0\right)=\begin{bmatrix}1/4&0\\0&0\end{bmatrix},$$

$$\Delta x_0\Delta g_0^T G_0=\begin{bmatrix}1/2\\0\end{bmatrix}(2,-1)=\begin{bmatrix}1&-1/2\\0&0\end{bmatrix},$$

$$G_0\Delta g_0\Delta x_0^T=\begin{bmatrix}2\\-1\end{bmatrix}\left(\frac{1}{2},0\right)=\begin{bmatrix}1&0\\-1/2&0\end{bmatrix},$$

因此

$$G_1=\begin{bmatrix}1&0\\0&1\end{bmatrix}+(1+5)\begin{bmatrix}1/4&0\\0&0\end{bmatrix}-\begin{bmatrix}1&-1/2\\0&0\end{bmatrix}-\begin{bmatrix}1&0\\-1/2&0\end{bmatrix}$$
$$=\begin{bmatrix}1/2&1/2\\1/2&1\end{bmatrix},$$

由此构造 BFGS 方向

$$d_1=-G_1\nabla f(x_1)=\begin{bmatrix}1/2\\1\end{bmatrix}.$$

从点 x_1 出发沿 d_1 进行一维搜索,得

$$\lambda_1=1,x_2=x_1+\lambda_1 d_1=\begin{bmatrix}1\\1\end{bmatrix}.$$

即知 $x_2=(1,1)^T$ 为问题(7.2.6)的最优解. □

7.5 最小二乘法

在实际应用中,会经常遇到目标函数为若干个函数平方和的最优化问题:

$$\min s(x)=\sum_{i=1}^{m}f_i^2(x),\tag{7.5.1}$$

其中 $x\in\mathbb{R}^n$,一般假设 $m\geqslant n$,这类问题称为最小二乘问题(least square problem).当每个 $f_i(x)$ 都是线性函数时,问题(7.5.1)称为线性最小二乘问题(least square linear problem),否则,称为非线性最小二乘问题(least square nonlinear problem).

由于最小二乘问题相对于一般无约束最优化问题而言具有特殊形式,因此除能运用本章前面介绍的一般求解方法外,还应有更为简便有效的方法.

7.5.1 线性最小二乘法

当 $f_i(x)$ 为线性函数时,即

$$f_i(\boldsymbol{x}) = \sum_{j=1}^{n} a_{ij}x_j - b_i, i = 1,2,\cdots,m,$$

问题(7.5.1)就成为线性最小二乘问题. 如令

$$\boldsymbol{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}, \boldsymbol{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix},$$

$$\boldsymbol{f}(\boldsymbol{x}) = (f_1(\boldsymbol{x}), f_2(\boldsymbol{x}), \cdots, f_m(\boldsymbol{x}))^{\mathrm{T}},$$

则

$$s(\boldsymbol{x}) = \boldsymbol{f}(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{f}(\boldsymbol{x}) = (\boldsymbol{Ax} - \boldsymbol{b})^{\mathrm{T}}(\boldsymbol{Ax} - \boldsymbol{b}) = \| \boldsymbol{Ax} - \boldsymbol{b} \|^2,$$

从而问题(7.5.1)可表示为

$$\min s(\boldsymbol{x}) = \| \boldsymbol{Ax} - \boldsymbol{b} \|^2. \tag{7.5.2}$$

因为

$$s(\boldsymbol{x}) = (\boldsymbol{Ax} - \boldsymbol{b})^{\mathrm{T}}(\boldsymbol{Ax} - \boldsymbol{b}) = \boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{Ax} - 2\boldsymbol{b}^{\mathrm{T}}\boldsymbol{Ax} + \boldsymbol{b}^{\mathrm{T}}\boldsymbol{b},$$

所以 $s(\boldsymbol{x})$ 为二次函数,且

$$\nabla s(\boldsymbol{x}) = 2\boldsymbol{A}^{\mathrm{T}}\boldsymbol{Ax} - 2\boldsymbol{A}^{\mathrm{T}}\boldsymbol{b}, \nabla^2 s(\boldsymbol{x}) = 2\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}.$$

显然,对一切 $\boldsymbol{y} \in \mathbb{R}^n$,均有

$$\boldsymbol{y}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{Ay} = \| \boldsymbol{Ay} \|^2 \geqslant 0,$$

即知,$\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}$ 是半正定矩阵,从而由定理 2.3.7 知,$s(\boldsymbol{x})$是$\mathbb{R}^n$ 上的凸函数.

定理 7.5.1 $\bar{\boldsymbol{x}} \in \mathbb{R}^n$ 是问题(7.5.2)的最优解的充要条件是 $\bar{\boldsymbol{x}}$ 满足方程组

$$\boldsymbol{A}^{\mathrm{T}}\boldsymbol{Ax} = \boldsymbol{A}^{\mathrm{T}}\boldsymbol{b}. \tag{7.5.3}$$

证明 因为 $s(\boldsymbol{x})$ 是在$\mathbb{R}^n$ 上具有一阶连续偏导数的凸函数,所以由推论 3.1.5 知,$\bar{\boldsymbol{x}}$ 是问题(7.5.2)的最优解的充要条件是$\nabla s(\bar{\boldsymbol{x}}) = \boldsymbol{0}$,即

$$\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}\bar{\boldsymbol{x}} = \boldsymbol{A}^{\mathrm{T}}\boldsymbol{b}.$$

□

容易知道,矩阵 $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}$ 正定的充要条件是对于一切非零向量 $\boldsymbol{y} \in \mathbb{R}^n$,有

$$\boldsymbol{y}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{Ay} = \| \boldsymbol{Ay} \|^2 > 0.$$

记 $\boldsymbol{y} = (y_1, y_2, \cdots, y_n)^{\mathrm{T}}, \boldsymbol{A} = (\boldsymbol{p}_1, \boldsymbol{p}_2, \cdots, \boldsymbol{p}_n)$,则上式等价于

$$\boldsymbol{Ay} = \sum_{j=1}^{n} \boldsymbol{p}_j y_j \neq \boldsymbol{0}. \tag{7.5.4}$$

而(7.5.4)式成立的充要条件是向量组 $\boldsymbol{p}_1, \boldsymbol{p}_2, \cdots, \boldsymbol{p}_n$ 线性无关,这又等价于 $\operatorname{rank} \boldsymbol{A} = n$. 从而 $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}$ 为正定矩阵的充要条件是 $\operatorname{rank} \boldsymbol{A} = n$,换句话说,$s(\boldsymbol{x})$为正定二次函数的充要条件是 $\operatorname{rank} \boldsymbol{A} = n$.

定理 7.5.2 若 $\operatorname{rank} \boldsymbol{A} = n$,则

$$\bar{\boldsymbol{x}} = (\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A})^{-1}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{b} \tag{7.5.5}$$

为问题(7.5.2)的惟一最优解.

证明 因 rank $A=n$,故由前述可知,$A^{\mathrm{T}}A$ 为正定矩阵,从而 $\bar{x}=(A^{\mathrm{T}}A)^{-1}A^{\mathrm{T}}b$ 为方程组(7.5.3)的惟一的解,于是,由定理 7.5.1 知,$\bar{x}$ 为问题(7.5.2)的惟一最优解. □

显然,方程组 $Ax=b$ 有解的充要条件是问题(7.5.2)的最优值$\min s(x)=0$.

例 7.5.1 求解线性最小二乘问题

$$\min \|Ax-b\|^2, \tag{7.5.6}$$

其中

$$A=\begin{bmatrix} 3 & 1 \\ 2 & -3 \\ -1 & 4 \end{bmatrix}, b=\begin{bmatrix} 2 \\ -3 \\ -1 \end{bmatrix}.$$

解 因为

$$A^{\mathrm{T}}A=\begin{bmatrix} 14 & -7 \\ -7 & 26 \end{bmatrix}, (A^{\mathrm{T}}A)^{-1}=\frac{1}{350}\begin{bmatrix} 26 & 7 \\ 7 & 14 \end{bmatrix},$$

所以,由公式(7.5.5),求得问题(7.5.6)的最优解

$$\bar{x}=\frac{1}{350}\begin{bmatrix} 26 & 7 \\ 7 & 14 \end{bmatrix}\begin{bmatrix} 3 & 2 & -1 \\ 1 & -3 & 4 \end{bmatrix}\begin{bmatrix} 2 \\ -3 \\ -1 \end{bmatrix}=\begin{bmatrix} 3/14 \\ 3/10 \end{bmatrix}.$$

由于问题(7.5.6)的最优值为

$$(A\bar{x}-b)^{\mathrm{T}}(A\bar{x}-b)=11.454\ 285\ 71\neq 0,$$

因此,方程组

$$\begin{cases} 3x_1+x_2=2, \\ 2x_1-3x_2=-3, \\ -x_1+4x_2=-1 \end{cases}$$

无解. □

7.5.2 Gauss - Newton 法

现在讨论非线性最小二乘问题

$$\min s(x)=\|f(x)\|^2, \tag{7.5.7}$$

其中 $x\in\mathbb{R}^n$,$f(x)=(f_1(x),f_2(x),\cdots,f_m(x))^{\mathrm{T}}$,$f_i(x)$不全是线性函数,且假定$f_i(x)$具有一阶连续偏导数,$i=1,2,\cdots,m$.

求解问题(7.5.7)的基本思想与 7.2 节中的 Newton 法类似:把 $f_i(x)$ 线性化,用线性最小二乘问题的解去逼近非线性最小二乘问题的解.

假设选定初始点 x_0 后经过 k 次迭代得到 x_k,把 $f_i(x)$ 在点 x_k 处展成一阶 Taylor 表达式(Newton 法是展成二阶 Taylor 表达式)

$$f_i(x)\approx f_i(x_k)+\nabla f_i(x_k)^{\mathrm{T}}(x-x_k). \tag{7.5.8}$$

根据$\boldsymbol{f}(\boldsymbol{x})$的 Jacobi 矩阵

$$\nabla \boldsymbol{f}(\boldsymbol{x}) = (\nabla f_1(\boldsymbol{x}), \nabla f_2(\boldsymbol{x}), \cdots, \nabla f_m(\boldsymbol{x}))^{\mathrm{T}}$$

可知,(7.5.8)式可写成

$$\boldsymbol{f}(\boldsymbol{x}) \approx \boldsymbol{f}(\boldsymbol{x}_k) + \nabla \boldsymbol{f}(\boldsymbol{x}_k)(\boldsymbol{x} - \boldsymbol{x}_k),$$

从而

$$s(\boldsymbol{x}) \approx \| \boldsymbol{f}(\boldsymbol{x}_k) + \nabla \boldsymbol{f}(\boldsymbol{x}_k)(\boldsymbol{x} - \boldsymbol{x}_k) \|^2.$$

因此,线性最小二乘问题

$$\min \| \nabla \boldsymbol{f}(\boldsymbol{x}_k)(\boldsymbol{x} - \boldsymbol{x}_k) + \boldsymbol{f}(\boldsymbol{x}_k) \|^2 \tag{7.5.9}$$

的最优解可以作为问题(7.5.7)的最优解的第 $k+1$ 次近似,记为 $\boldsymbol{x}_{k+1}$.

根据定理 7.5.1,问题(7.5.9)的最优解 $\boldsymbol{x}_{k+1}$满足方程组

$$\nabla \boldsymbol{f}(\boldsymbol{x}_k)^{\mathrm{T}} \nabla \boldsymbol{f}(\boldsymbol{x}_k)(\boldsymbol{x} - \boldsymbol{x}_k) = -\nabla \boldsymbol{f}(\boldsymbol{x}_k)^{\mathrm{T}} \boldsymbol{f}(\boldsymbol{x}_k).$$

再假设矩阵$\nabla \boldsymbol{f}(\boldsymbol{x}_k)$是列满秩的,则由定理 7.5.2 可知,上述方程有惟一的解

$$\boldsymbol{x}_{k+1} = \boldsymbol{x}_k - [\nabla \boldsymbol{f}(\boldsymbol{x}_k)^{\mathrm{T}} \nabla \boldsymbol{f}(\boldsymbol{x}_k)]^{-1} \nabla \boldsymbol{f}(\boldsymbol{x}_k)^{\mathrm{T}} \boldsymbol{f}(\boldsymbol{x}_k).$$

上式称为 Gauss - Newton 公式,向量

$$\boldsymbol{d}_k = -[\nabla \boldsymbol{f}(\boldsymbol{x}_k)^{\mathrm{T}} \nabla \boldsymbol{f}(\boldsymbol{x}_k)]^{-1} \nabla \boldsymbol{f}(\boldsymbol{x}_k)^{\mathrm{T}} \boldsymbol{f}(\boldsymbol{x}_k) \tag{7.5.10}$$

称为 Gauss - Newton 方向. 由 Gauss - Newton 公式确定的迭代算法称为 Gauss - Newton 法.

因为$\nabla \boldsymbol{f}(\boldsymbol{x}_k)^{\mathrm{T}} \nabla \boldsymbol{f}(\boldsymbol{x}_k)$正定,所以$[\nabla \boldsymbol{f}(\boldsymbol{x}_k)^{\mathrm{T}} \nabla \boldsymbol{f}(\boldsymbol{x}_k)]^{-1}$是正定矩阵. 于是由

$$\nabla s(\boldsymbol{x}) = 2 \sum_{i=1}^{m} \nabla f_i(\boldsymbol{x}) f_i(\boldsymbol{x}) = 2 \nabla \boldsymbol{f}(\boldsymbol{x})^{\mathrm{T}} \boldsymbol{f}(\boldsymbol{x}),$$

及(7.5.10)式可得

$$\nabla s(\boldsymbol{x}_k)^{\mathrm{T}} \boldsymbol{d}_k = -2[\nabla \boldsymbol{f}(\boldsymbol{x}_k)^{\mathrm{T}} \boldsymbol{f}(\boldsymbol{x}_k)]^{\mathrm{T}} [\nabla \boldsymbol{f}(\boldsymbol{x}_k)^{\mathrm{T}} \nabla \boldsymbol{f}(\boldsymbol{x}_k)]^{-1} [\nabla \boldsymbol{f}(\boldsymbol{x}_k)^{\mathrm{T}} \boldsymbol{f}(\boldsymbol{x}_k)] < 0,$$

由此即知,Gauss - Newton 方向是 $s(\boldsymbol{x})$在点 $\boldsymbol{x}_k$ 处的下降方向.

同 Newton 法类似,可以证明,当初始点 $\boldsymbol{x}_0$ 充分接近问题(7.5.7)的极小点 $\bar{\boldsymbol{x}}$ 时,则 Gauss - Newton 法是二阶收敛的.

为了保证在初始点 x_0 远离问题(7.5.7)的极小点 $\bar{x}$ 时,算法仍然收敛,与阻尼 Newton 法一样,在求出 $\boldsymbol{x}_k$ 和 $\boldsymbol{d}_k$ 之后,不直接用 $\boldsymbol{x}_k + \boldsymbol{d}_k$ 作为 $k+1$ 次近似,而是从 $\boldsymbol{x}_k$ 出发沿 $\boldsymbol{d}_k$ 进行最优一维搜索

$$s(\boldsymbol{x}_k + \lambda_k \boldsymbol{d}_k) = \min_{\lambda \geqslant 0} s(\boldsymbol{x}_k + \lambda \boldsymbol{d}_k),$$

并取 $\boldsymbol{x}_{k+1} = \boldsymbol{x}_k + \lambda_k \boldsymbol{d}_k$ 作为第 $k+1$ 次近似.

算法 7-8(阻尼 Gauss - Newton 法)

Step 1　选取初始数据. 选取初始点 $\boldsymbol{x}_0$,给定允许误差 $\varepsilon > 0$,令 $k = 0$.

Step 2　检查是否满足终止准则. 计算 $\boldsymbol{f}(\boldsymbol{x}_k)$及 Jacobi 矩阵$\nabla \boldsymbol{f}(\boldsymbol{x}_k)$,若 $\| \nabla \boldsymbol{f}(\boldsymbol{x}_k)^{\mathrm{T}} \boldsymbol{f}(\boldsymbol{x}_k) \| < \varepsilon$,迭代终止,$\boldsymbol{x}_k$ 为(7.5.7)的近似极小点;否则,转 Step 3.

Step 3　构造 Gauss－Newton 方向. 计算$[\nabla \boldsymbol{f}(\boldsymbol{x}_k)^{\mathrm{T}} \nabla \boldsymbol{f}(\boldsymbol{x}_k)]^{-1}$,取

$$\boldsymbol{d}_k = -[\nabla \boldsymbol{f}(\boldsymbol{x}_k)^{\mathrm{T}} \nabla \boldsymbol{f}(\boldsymbol{x}_k)]^{-1} \nabla \boldsymbol{f}(\boldsymbol{x}_k)^{\mathrm{T}} \boldsymbol{f}(\boldsymbol{x}_k).$$

Step 4　进行一维搜索. 求 λ_k 和 $\boldsymbol{x}_{k+1}$,使得

$$s(\boldsymbol{x}_k + \lambda_k \boldsymbol{d}_k) = \min_{\lambda \geqslant 0} s(\boldsymbol{x}_k + \lambda \boldsymbol{d}_k),$$

$$\boldsymbol{x}_{k+1} = \boldsymbol{x}_k + \lambda_k \boldsymbol{d}_k.$$

令 $k := k+1$,返回 Step 2.

同阻尼 Newton 法一样,阻尼 Gauss－Newton 法具有全局收敛性.

7.5.3 Levenberg－Marquardt 法

在 Gauss－Newton 法中,我们要求$\nabla \boldsymbol{f}(\boldsymbol{x}_k)$是列满秩的. 但遗憾的是,常常出现$\nabla \boldsymbol{f}(\boldsymbol{x}_k)$不是列满秩的情形. 此时,迭代将无法进行下去. 为此,Levenberg (1944)和 Marquardt(1963)对 Gauss－Newton 法作了如下修正:把一个正定矩阵 $\alpha_k \boldsymbol{I}_n$ 加到$\nabla \boldsymbol{f}(\boldsymbol{x}_k)^{\mathrm{T}} \nabla \boldsymbol{f}(\boldsymbol{x}_k)$上去($\alpha_k > 0$ 为参数,$\boldsymbol{I}_n$ 为 n 阶单位矩阵),使之成为正定矩阵,然后令

$$\boldsymbol{d}_k = -[\nabla \boldsymbol{f}(\boldsymbol{x}_k)^{\mathrm{T}} \nabla \boldsymbol{f}(\boldsymbol{x}_k) + \alpha_k \boldsymbol{I}_n]^{-1} \nabla \boldsymbol{f}(\boldsymbol{x}_k)^{\mathrm{T}} \boldsymbol{f}(\boldsymbol{x}_k), \qquad (7.5.11)$$

$$\boldsymbol{x}_{k+1} = \boldsymbol{x}_k + \boldsymbol{d}_k, \qquad (7.5.12)$$

由(7.5.11)和(7.5.12)两式确定的迭代算法称为 Levenberg－Marquardt 法,简称为最小二乘 LM 法. 由(7.5.11)式确定的 $\boldsymbol{d}_k$ 称为 LM 方向.

容易知道,当$\nabla \boldsymbol{f}(\boldsymbol{x}_k)^{\mathrm{T}} \nabla \boldsymbol{f}(\boldsymbol{x}_k) + \alpha_k \boldsymbol{I}_n$ 为正定矩阵时,由(7.5.11)式确定的 $\boldsymbol{d}_k$ 是 $s(\boldsymbol{x})$在点 $\boldsymbol{x}_k$ 处的下降方向.

定理 7.5.3　设 $\alpha_k > 0$,$\boldsymbol{d}_k$ 由(7.5.11)式确定,则 $\boldsymbol{d}_k$ 是最优化问题

$$\begin{cases} \min \| \nabla \boldsymbol{f}(\boldsymbol{x}_k) \boldsymbol{y} + \boldsymbol{f}(\boldsymbol{x}_k) \|^2; \\ \text{s.t. } \| \boldsymbol{y} \| \leqslant \| \boldsymbol{d}_k \| \end{cases} \qquad (7.5.13)$$

的全局极小点;并且,当 α_k 适当大时,有

$$s(\boldsymbol{x}_k + \boldsymbol{d}_k) < s(\boldsymbol{x}_k); \qquad (7.5.14)$$

当 α_k 充分大时,方向 $\boldsymbol{d}_k$ 与方向 $-\nabla s(\boldsymbol{x}_k)$充分接近.

证明　(1) 显然 $\boldsymbol{d}_k$ 是问题(7.5.13)的可行点. 记

$$\varphi(\boldsymbol{y}) = \| \nabla \boldsymbol{f}(\boldsymbol{x}_k) \boldsymbol{y} + \boldsymbol{f}(\boldsymbol{x}_k) \|^2,$$

则由(7.5.11)式可知

$$\begin{aligned} \varphi(\boldsymbol{y}) &= \boldsymbol{y}^{\mathrm{T}} \nabla \boldsymbol{f}(\boldsymbol{x}_k)^{\mathrm{T}} \nabla \boldsymbol{f}(\boldsymbol{x}_k) \boldsymbol{y} + 2\boldsymbol{y}^{\mathrm{T}} \nabla \boldsymbol{f}(\boldsymbol{x}_k)^{\mathrm{T}} \boldsymbol{f}(\boldsymbol{x}_k) + \boldsymbol{f}(\boldsymbol{x}_k)^{\mathrm{T}} \boldsymbol{f}(\boldsymbol{x}_k) \\ &= \boldsymbol{y}^{\mathrm{T}} \nabla \boldsymbol{f}(\boldsymbol{x}_k)^{\mathrm{T}} \nabla \boldsymbol{f}(\boldsymbol{x}_k) \boldsymbol{y} - 2\boldsymbol{y}^{\mathrm{T}} [\nabla \boldsymbol{f}(\boldsymbol{x}_k)^{\mathrm{T}} \nabla \boldsymbol{f}(\boldsymbol{x}_k) + \alpha_k \boldsymbol{I}_n] \boldsymbol{d}_k + \boldsymbol{f}(\boldsymbol{x}_k)^{\mathrm{T}} \boldsymbol{f}(\boldsymbol{x}_k), \end{aligned}$$

从而

$$\varphi(\boldsymbol{d}_k) = -\boldsymbol{d}_k^{\mathrm{T}} \nabla \boldsymbol{f}(\boldsymbol{x}_k)^{\mathrm{T}} \nabla \boldsymbol{f}(\boldsymbol{x}_k) \boldsymbol{d}_k - 2\alpha_k \boldsymbol{d}_k^{\mathrm{T}} \boldsymbol{d}_k + \boldsymbol{f}(\boldsymbol{x}_k)^{\mathrm{T}} \boldsymbol{f}(\boldsymbol{x}_k).$$

于是对于满足 $\| \boldsymbol{y} \| \leqslant \| \boldsymbol{d}_k \|$ 的任意 $\boldsymbol{y}$,利用 Cauchy 不等式,并注意到 $\alpha_k > 0$,

$\nabla f(\boldsymbol{x}_k)^{\mathrm{T}} \nabla f(\boldsymbol{x}_k)$是半正定矩阵,可得

$$
\begin{aligned}
\varphi(\boldsymbol{y})-\varphi(\boldsymbol{d}_k) &= (\boldsymbol{d}_k-\boldsymbol{y})^{\mathrm{T}} \nabla f(\boldsymbol{x}_k)^{\mathrm{T}} \nabla f(\boldsymbol{x}_k)(\boldsymbol{d}_k-\boldsymbol{y}) + 2\alpha_k(\boldsymbol{d}_k-\boldsymbol{y})^{\mathrm{T}}\boldsymbol{d}_k \\
&\geqslant (\boldsymbol{d}_k-\boldsymbol{y})^{\mathrm{T}} \nabla f(\boldsymbol{x}_k)^{\mathrm{T}} \nabla f(\boldsymbol{x}_k)(\boldsymbol{d}_k-\boldsymbol{y}) + 2\alpha_k(\|\boldsymbol{d}_k\|-\|\boldsymbol{y}\|)\|\boldsymbol{d}_k\| \\
&\geqslant 0,
\end{aligned}
$$

这说明 $\boldsymbol{d}_k$ 是问题(7.5.13)的全局极小点.

(2) 设半正定矩阵的$\nabla f(\boldsymbol{x}_k)^{\mathrm{T}} \nabla f(\boldsymbol{x}_k)$特征值为 $\mu_1,\mu_2,\cdots,\mu_n$,则 $\mu_j\geqslant 0$, $j=1,2,\cdots,n$,且存在正交矩阵 $\boldsymbol{P}$,使

$$
\boldsymbol{P}^{\mathrm{T}}[\nabla f(\boldsymbol{x}_k)^{\mathrm{T}} f(\boldsymbol{x}_k)]\boldsymbol{P} = \boldsymbol{D} = \mathrm{diag}(\mu_1,\mu_2,\cdots,\mu_n).
$$

由 $\boldsymbol{P}^{\mathrm{T}}\boldsymbol{P}=\boldsymbol{P}\boldsymbol{P}^{\mathrm{T}}=\boldsymbol{I}_n$ 和$\nabla f(\boldsymbol{x}_k)^{\mathrm{T}} \nabla f(\boldsymbol{x}_k)=\boldsymbol{P}\boldsymbol{D}\boldsymbol{P}^{\mathrm{T}}$ 及(7.5.11)式有

$$
\boldsymbol{d}_k = -\boldsymbol{P}(\boldsymbol{D}+\alpha_k\boldsymbol{I}_n)^{-1}\boldsymbol{P}^{\mathrm{T}} \nabla f(\boldsymbol{x}_k)^{\mathrm{T}} f(\boldsymbol{x}_k). \tag{7.5.15}
$$

记

$$
\boldsymbol{z} = (z_1,z_2,\cdots,z_n)^{\mathrm{T}} = \boldsymbol{P}^{\mathrm{T}} \nabla f(\boldsymbol{x}_k)^{\mathrm{T}} f(\boldsymbol{x}_k), \tag{7.5.16}
$$

则由(7.5.15)式可得

$$
\|\boldsymbol{d}_k\|^2 = \boldsymbol{z}^{\mathrm{T}}[(\boldsymbol{D}+\alpha_k\boldsymbol{I}_n)^{-1}]^2\boldsymbol{z} = \sum_{i=1}^{n}\frac{z_i^2}{(\mu_i+\alpha_k)^2}, \tag{7.5.17}
$$

由此显然有

$$
\lim_{\alpha_k\to\infty}\|\boldsymbol{d}_k\|^2 = 0. \tag{7.5.18}
$$

注意到(7.5.11)式确定的方向 $\boldsymbol{d}_k$ 是 $s(\boldsymbol{x})$ 在点 $\boldsymbol{x}_k$ 处的下降方向,当 α_k 适当大时,(7.5.18)式表明 $\boldsymbol{x}_k+\boldsymbol{d}_k$ 是从点 $\boldsymbol{x}_k$ 出发沿下降方向作很小的移动,故(7.5.14)式成立.

(3) 记 $\boldsymbol{d}_k$ 与 $-\nabla s(\boldsymbol{x}_k)$的夹角为 φ,则

$$
\cos\varphi = \frac{-\boldsymbol{d}_k^{\mathrm{T}} \nabla s(\boldsymbol{x}_k)}{\|\boldsymbol{d}_k\| \, \|\nabla s(\boldsymbol{x}_k)\|}.
$$

由(7.5.16)和(7.5.15)两式得

$$
\nabla s(\boldsymbol{x}_k) = 2\nabla f(\boldsymbol{x}_k)^{\mathrm{T}} f(\boldsymbol{x}_k) = 2\boldsymbol{P}\boldsymbol{z},
$$

$$
\boldsymbol{d}_k = -\boldsymbol{P}(\boldsymbol{D}+\alpha_k\boldsymbol{I}_n)^{-1}\boldsymbol{z},
$$

从而根据(7.5.17)式可得

$$
\begin{aligned}
\cos\varphi &= \frac{2\boldsymbol{z}^{\mathrm{T}}[(\boldsymbol{D}+\alpha_k\boldsymbol{I}_n)^{-1}]^{\mathrm{T}}\boldsymbol{P}^{\mathrm{T}}\boldsymbol{P}\boldsymbol{z}}{\|\boldsymbol{d}_k\|\cdot 2\|\boldsymbol{z}\|} \\
&= \left[\sum_{i=1}^{n}\frac{z_i^2}{\mu_i+\alpha_k}\right]\left[\sum_{i=1}^{n}\frac{z_i^2}{(\mu_i+\alpha_k)^2}\right]^{-\frac{1}{2}}\left[\sum_{i=1}^{n}z_i^2\right]^{-\frac{1}{2}},
\end{aligned}
$$

于是

$$
\lim_{\alpha_k\to\infty}\cos\varphi = 1,
$$

这表明,当 α_k 充分大时,方向 $\boldsymbol{d}_k$ 与方向 $-\nabla s(\boldsymbol{x}_k)$接近. □

我们知道，Gauss - Newton 方向是线性最小二乘问题

$$\min \| \nabla \boldsymbol{f}(\boldsymbol{x}_k)\boldsymbol{y} + \boldsymbol{f}(\boldsymbol{x}_k) \|^2$$

的最优解，LM 方向 $\boldsymbol{d}_k$ 虽然依赖于参数 α_k，但定理 7.5.3 表明，$\boldsymbol{d}_k$ 也是相应的线性最小二乘问题在范数限制下的最优解. 在最小二乘 LM 法中，当 $\alpha_k = 0$ 时，由 (7.5.11) 式确定的 $\boldsymbol{d}_k$ 就是 Gauss - Newton 方向. 根据定理 7.5.3 知，随着 α_k 的增大，LM 方向逐步向 $-\nabla s(\boldsymbol{x}_k)$ 偏移，当 α_k 充分大时便充分接近 $s(\boldsymbol{x})$ 在点 $\boldsymbol{x}_k$ 处的最速下降方向. 因此，在最小二乘 LM 法的迭代中要限制 α_k 值的增大，否则，会减慢算法的收敛速度. 另一方面，如果 α_k 太小，则由定理 7.5.3 知不能保证在迭代过程中使目标函数值下降. 所以，如何确定参数 α_k 是 LM 法的一个重要问题.

算法 7 - 9(最小二乘 LM 法)

Step 1　选取初始数据. 选取初始点 x_0，给定初始参数 $\alpha_0 > 0$，放大因子 $\beta > 1$ 及允许误差 $\varepsilon > 0$.

Step 2　求初始目标函数值. 计算 $\boldsymbol{f}(\boldsymbol{x}_0)$ 及 $s(\boldsymbol{x}_0)$，令 $k = 0$.

Step 3　求 Jacobi 矩阵. 计算 $\nabla \boldsymbol{f}(\boldsymbol{x}_k)$.

Step 4　检查是否满足终止准则. 若

$$\| \nabla \boldsymbol{f}(\boldsymbol{x}_k)^{\mathrm{T}} \boldsymbol{f}(\boldsymbol{x}_k) \| < \varepsilon,$$

迭代终止，$\boldsymbol{x}_k$ 为问题 (7.5.7) 的近似最优解；否则，转 Step 5.

Step 5　构造 LM 方向. 计算 $[\nabla \boldsymbol{f}(\boldsymbol{x}_k)^{\mathrm{T}} \nabla \boldsymbol{f}(\boldsymbol{x}_k) + \alpha_k \boldsymbol{I}_n]^{-1}$，令

$$\boldsymbol{d}_k = -[\nabla \boldsymbol{f}(\boldsymbol{x}_k)^{\mathrm{T}} \nabla \boldsymbol{f}(\boldsymbol{x}_k) + \alpha_k \boldsymbol{I}_n]^{-1} \nabla \boldsymbol{f}(\boldsymbol{x}_k)^{\mathrm{T}} \boldsymbol{f}(\boldsymbol{x}_k).$$

Step 6　检查目标函数是否下降. 计算 $\boldsymbol{f}(\boldsymbol{x}_k + \boldsymbol{d}_k)$ 及 $s(\boldsymbol{x}_k + \boldsymbol{d}_k)$，若 $s(\boldsymbol{x}_k + \boldsymbol{d}_k) < s(\boldsymbol{x}_k)$，转 Step 8；否则，转 Step 7.

Step 7　放大参数. 令 $\alpha_k := \beta \alpha_k$，返回 Step 5.

Step 8　缩小参数. 令 $\boldsymbol{x}_{k+1} = \boldsymbol{x}_k + \boldsymbol{d}_k$，$\alpha_{k+1} = \alpha_k / \beta$，$k := k + 1$，返回 Step 3.

初始参数 α_0 和放大因子 β 应取适当数值，例如，根据经验可取 $\alpha_0 = 0.01$，$\beta = 10$.

可以证明，最小二乘 LM 法具有全局收敛性，详细证明参见文献[16,50].

7.6 信赖域法

我们知道 Newton 法存在两点不足：一是不具备全局收敛性，二是要求目标函数的 Hesse 矩阵正定. 7.2 节中采用最优一维搜索的阻尼 Newton 法克服了第一点不足；为了克服第二点不足，本节介绍一种新的方法——信赖域法.

假设问题 (7.1.1) 的目标函数 f 具有二阶连续偏导数，第 k 次迭代点为 $\boldsymbol{x}_k$，根据二阶 Taylor 展开式，有

$$f(\boldsymbol{x}_k+\boldsymbol{y})\approx\varphi_k(\boldsymbol{y})=f(\boldsymbol{x}_k)+\nabla f(\boldsymbol{x}_k)^{\mathrm{T}}\boldsymbol{y}+\frac{1}{2}\boldsymbol{y}^{\mathrm{T}}\nabla^2 f(\boldsymbol{x}_k)\boldsymbol{y}.$$

如果$\nabla^2 f(\boldsymbol{x}_k)$正定,那么可以用 $\varphi_k(\boldsymbol{y})$的极小点 $\boldsymbol{d}_k$ 来修正 $\boldsymbol{x}_k$,得到 $\boldsymbol{x}_{k+1}=\boldsymbol{x}_k+\boldsymbol{d}_k$,这就是 Newton 迭代的思想;否则 $\varphi_k(\boldsymbol{y})$可能没有极小点,从而使得 Newton 迭代无法进行下去,并且只有当 $\|\boldsymbol{y}\|$ 很小时 $\varphi_k(\boldsymbol{y})$才能较好地近似于$f(\boldsymbol{x}_k+\boldsymbol{y})$,因此,我们用 $\varphi_k(\boldsymbol{y})$在

$$D(h_k)=\{\boldsymbol{y}\mid \|\boldsymbol{y}\|\leqslant h_k, h_k>0\}$$

上的极小点 $\boldsymbol{d}_k$ 来修正 $\boldsymbol{x}_k$,其中参数 h_k 的选取使得$f(\boldsymbol{x})$从 $\boldsymbol{x}_k$ 到 $\boldsymbol{x}_k+\boldsymbol{d}_k$ 的实际下降量

$$\Delta f_k=f(\boldsymbol{x}_k)-f(\boldsymbol{x}_k+\boldsymbol{d}_k)>0,$$

此时称区域 $D(h_k)$为信赖域(trust region),称参数 h_k 为信赖域半径. 这样得到信赖域子问题

$$\begin{cases}\min\varphi_k(\boldsymbol{y})=f(\boldsymbol{x}_k)+\nabla f(\boldsymbol{x}_k)^{\mathrm{T}}\boldsymbol{y}+\frac{1}{2}\boldsymbol{y}^{\mathrm{T}}\nabla^2 f(\boldsymbol{x}_k)\boldsymbol{y};\\ \text{s.t. } \|\boldsymbol{y}\|\leqslant h_k,\end{cases}\tag{7.6.1}$$

这个最优化问题的目标函数是二次函数,约束是一个非线性不等式.

7.6.1 信赖域半径的确定

首先要解决的是信赖域半径的存在性问题.

定理 7.6.1 若$\nabla f(\boldsymbol{x}_k)\neq\boldsymbol{0}$,则存在 $h_k>0$,使得问题(7.6.1)的全局极小点 $\boldsymbol{d}_k$ 满足$f(\boldsymbol{x}_k+\boldsymbol{d}_k)<f(\boldsymbol{x}_k)$,即 $\Delta f_k>0$.

证明 反证法. 假设结论不成立,令 $h_k=\frac{1}{k}(k=1,2,\cdots)$,则问题(7.6.1)的全局极小点 $\boldsymbol{d}_k$,使得

$$f(\boldsymbol{x}_k+\boldsymbol{d}_k)\geqslant f(\boldsymbol{x}_k), k=1,2,\cdots.\tag{7.6.2}$$

若 $\boldsymbol{d}_k=\boldsymbol{0}$,则由 $\boldsymbol{d}_k$ 是问题(7.6.1)的全局极小点可知,$\boldsymbol{d}_k$ 也是函数 $\varphi_k(\boldsymbol{y})$的一个局部极小点,从而由定理 3.1.1 可得

$$\nabla f(\boldsymbol{x}_k)=\nabla\varphi_k(\boldsymbol{0})=\boldsymbol{0},$$

与条件矛盾. 因此 $\boldsymbol{d}_k\neq\boldsymbol{0}$,即 $0<\|\boldsymbol{d}_k\|\leqslant h_k$, $\|\boldsymbol{d}_k\|\to 0(k\to\infty)$.

记 $\boldsymbol{p}_k=\dfrac{\boldsymbol{d}_k}{\|\boldsymbol{d}_k\|}$,则$\{\boldsymbol{p}_k\}$是有界点列,于是存在收敛的子列$\{\boldsymbol{p}_{k_m}\}$. 设$\lim\limits_{m\to\infty}\boldsymbol{p}_{k_m}=\bar{\boldsymbol{p}}$,显然 $\|\bar{\boldsymbol{p}}\|=1$.

根据一阶 Taylor 展开式,有

$$f(\boldsymbol{x}_k+\boldsymbol{d}_{k_m})=f(\boldsymbol{x}_k)+\nabla f(\boldsymbol{x}_k)^{\mathrm{T}}\boldsymbol{d}_{k_m}+o(\|\boldsymbol{d}_{k_m}\|), m=1,2,\cdots,$$

所以由(7.6.2)式,有

$$\nabla f(\boldsymbol{x}_k)^{\mathrm{T}}\boldsymbol{d}_{k_m}+o(\|\boldsymbol{d}_{k_m}\|)\geqslant 0, m=1,2,\cdots,$$

上式两边同除以 $\|d_{k_m}\|$，再令 $m\to\infty$，得

$$\nabla f(x_k)^{\mathrm{T}}\bar{p}\geqslant 0. \tag{7.6.3}$$

另一方面，因 d_k 是问题(7.6.1)的全局极小点，故必有

$$\varphi_k(d_{k_m})\leqslant\varphi_k\left(-\|d_{k_m}\|\frac{\nabla f(x_k)}{\|\nabla f(x_k)\|}\right),m=1,2,\cdots,$$

即

$$\|d_{k_m}\|\nabla f(x_k)^{\mathrm{T}}p_{k_m}+\frac{\|d_{k_m}\|^2}{2}p_{k_m}^{\mathrm{T}}\nabla^2 f(x_k)p_{k_m}$$

$$\leqslant-\|d_{k_m}\|\ \|\nabla f(x_k)\|+\frac{\|d_{k_m}\|^2}{2\|\nabla f(x_k)\|^2}\nabla f(x_k)^{\mathrm{T}}\nabla^2 f(x_k)\nabla f(x_k),$$

$$m=1,2,\cdots,$$

上式两边同除以 $\|d_{k_m}\|$，再令 $m\to\infty$，得

$$\nabla f(x_k)^{\mathrm{T}}\bar{p}\leqslant-\|\nabla f(x_k)\|<0,$$

此与(7.6.3)式矛盾. □

这个定理指出信赖域半径是存在的. 下面来讨论信赖域半径的选取方法.

令估计下降量

$$\Delta\varphi_k=f(x_k)-\varphi_k(d_k),$$

易知，当 $\nabla f(x_k)\neq\mathbf{0}$ 时，有 $\Delta\varphi_k>0$. 这是因为，若 $\Delta\varphi_k\leqslant 0$，则

$$\varphi_k(\mathbf{0})=f(x_k)\leqslant\varphi_k(d_k),$$

而 d_k 是问题(7.6.1)的全局极小点，即

$$\varphi_k(d_k)\leqslant\varphi_k(y),\ \forall\ \|y\|\leqslant\|d_k\|,$$

所以，$y=\mathbf{0}$ 是函数 $\varphi_k(y)$ 的一个局部极小点，从而由定理 3.1.1 有

$$\nabla f(x_k)=\nabla\varphi_k(\mathbf{0})=\mathbf{0}.$$

利用实际下降量与估计下降量的比值

$$r_k=\frac{\Delta f_k}{\Delta\varphi_k},$$

来衡量 $\varphi_k(d_k)$ 与 $f(x_k+d_k)$ 的近似程度. 当 r_k 接近于 1 时，表明近似程度好，取 $x_{k+1}=x_k+d_k$，应当使 h_k 的值不变或增大；当 r_k 接近 0 或小于 0 时，表明近似程度差. 取 $x_{k+1}=x_k$，应当减小 h_k 的值.

基于以上分析，可以给出信赖域的一个模式算法.

算法 7－10(信赖域法)

Step 1 选取初始数据. 选取初始点 x_0，给定允许误差 $\varepsilon>0$，初始信赖域半径 $h_0>0,0<\beta<\gamma<1,0<\theta<1,\rho>1$，令 $k=0$.

Step 2 检查是否满足终止准则，计算 $\nabla f(x_k)$ 和 $\nabla^2 f(x_k)$，若 $\|\nabla f(x_k)\|<\varepsilon$，迭代终止，$x_k$ 为问题(7.1.1)的近似最优解.

Step 3　求解二次规划问题(7.6.1),得到解 $\boldsymbol{d}_k$.

Step 4　计算 $f(\boldsymbol{x}_k+\boldsymbol{d}_k)$ 和 r_k 的值.

Step 5　若 $r_k<\beta$,则令 $h_{k+1}=\theta\|\boldsymbol{d}_k\|$;若 $r_k>\gamma$,且 $\|\boldsymbol{d}_k\|=h_k$,则令 $h_{k+1}=\rho h_k$;否则,令 $h_{k+1}=h_k$.

Step 6　若 $r_k\leqslant 0$,令 $\boldsymbol{x}_{x+1}=\boldsymbol{x}_k$;否则,令 $\boldsymbol{x}_{k+1}=\boldsymbol{x}_k+\boldsymbol{d}_k$.

Step 7　令 $k:=k+1$,转 Step 2.

根据经验,可选取信赖域法中参数 $h_0=\|\nabla f(\boldsymbol{x}_0)\|$,$\beta=0.25$,$\gamma=0.75$,$\theta=0.25$,$\rho=2$.

在适当的条件下,信赖域法具有全局收敛性,并且是二阶收敛的,参见文献[50].

7.6.2　信赖域子问题的求解

显然,问题(7.5.13)是一个信赖域子问题,所以根据定理 7.5.3,最小二乘 LM 法也是一种信赖域法. 因此,我们可以把最小二乘 LM 法的思想应用于求解信赖域子问题:对矩阵 $\nabla^2 f(\boldsymbol{x}_k)$,加上一个正定矩阵 $\alpha_k\boldsymbol{I}_n(\alpha_k>0)$,使之成为正定矩阵 $\nabla^2 f(\boldsymbol{x}_k)+\alpha_k\boldsymbol{I}_n$,从而弥补了 $\nabla^2 f(\boldsymbol{x}_k)$ 可能不正定的缺陷. 这就是信赖域 LM 法的基本思想.

定理 7.6.2　设 $\boldsymbol{d}_k$ 是问题(7.6.1)的可行点,则 $\boldsymbol{d}_k$ 为问题(7.6.1)的全局极小点的充要条件是存在数 $\alpha_k\geqslant 0$,使得

$$\begin{cases}[\nabla^2 f(\boldsymbol{x}_k)+\alpha_k\boldsymbol{I}_n]\boldsymbol{d}_k=-\nabla f(\boldsymbol{x}_k),\\ \alpha_k(h_k^2-\|\boldsymbol{d}_k\|^2)=0,\end{cases}\tag{7.6.4}$$

并且 $\nabla^2 f(\boldsymbol{x}_k)+\alpha_k\boldsymbol{I}_n$ 是半正定矩阵.

证明　必要性. 设 $\boldsymbol{d}_k$ 为问题(7.6.1)的全局极小点,则 K-T 条件成立,即(7.6.4)式成立. 再证 $\nabla^2 f(\boldsymbol{x}_k)+\alpha_k\boldsymbol{I}_n$ 是半正定矩阵.

当 $\|\boldsymbol{d}_k\|<h_k$ 时,$\alpha_k=0$,且 $\boldsymbol{d}_k$ 是函数 $\varphi_k(\boldsymbol{y})$ 的局部极小点,从而定理 3.1.2 可知,$\nabla^2\varphi_k(\boldsymbol{d}_k)=\nabla^2 f(\boldsymbol{x}_k)$ 是半正定矩阵,即 $\nabla^2 f(\boldsymbol{x}_k)+\alpha_k\boldsymbol{I}_n$ 是半正定矩阵.

当 $\|\boldsymbol{d}_k\|=h_k$ 时,由 $h_k>0$ 知 $\boldsymbol{d}_k\neq\boldsymbol{0}$. 根据(7.6.4)式中的第 1 式,对于一切 $\boldsymbol{y}\in\mathbb{R}^n$,有

$$\begin{aligned}\varphi_k(\boldsymbol{y})-\varphi_k(\boldsymbol{d}_k)&=\nabla f(\boldsymbol{x}_k)^{\mathrm{T}}(\boldsymbol{y}-\boldsymbol{d}_k)+\frac{1}{2}\boldsymbol{y}^{\mathrm{T}}\nabla^2 f(\boldsymbol{x}_k)\boldsymbol{y}-\frac{1}{2}\boldsymbol{d}_k^{\mathrm{T}}\nabla^2 f(\boldsymbol{x}_k)\boldsymbol{d}_k\\&=\frac{1}{2}(\boldsymbol{y}-\boldsymbol{d}_k)^{\mathrm{T}}(\nabla^2 f(\boldsymbol{x}_k)+\alpha_k\boldsymbol{I}_n)(\boldsymbol{y}-\boldsymbol{d}_k)+\frac{\alpha_k}{2}(\|\boldsymbol{d}_k\|^2-\|\boldsymbol{y}\|^2).\end{aligned}\tag{7.6.5}$$

因 $\boldsymbol{d}_k$ 为问题(7.6.1)的全局极小点,故

$$\varphi_k(\boldsymbol{y})\geqslant\varphi_k(\boldsymbol{d}_k),\ \forall\ \|\boldsymbol{y}\|\leqslant\|\boldsymbol{d}_k\|,$$

于是由(7.6.5)式知

$$\frac{1}{2}(\boldsymbol{y}-\boldsymbol{d}_k)^{\mathrm{T}}(\nabla^2 f(\boldsymbol{x}_k)+\alpha_k\boldsymbol{I}_n)(\boldsymbol{y}-\boldsymbol{d}_k)\geqslant 0,\ \forall\ \|\boldsymbol{y}\|=\|\boldsymbol{d}_k\|. \tag{7.6.6}$$

对于任意 $\boldsymbol{x}\in\mathbb{R}^n$,若 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{d}_k\neq 0$,取

$$t=-\frac{\|\boldsymbol{x}\|^2}{2\boldsymbol{x}^{\mathrm{T}}\boldsymbol{d}_k},\boldsymbol{y}=\frac{1}{\|\boldsymbol{x}\|^2}(\boldsymbol{x}^{\mathrm{T}}\boldsymbol{x}\boldsymbol{d}_k-2\boldsymbol{x}^{\mathrm{T}}\boldsymbol{d}_k\boldsymbol{x}),$$

则有

$$\boldsymbol{x}=t(\boldsymbol{y}-\boldsymbol{d}_k),\ \|\boldsymbol{y}\|=\|\boldsymbol{d}_k\|,$$

从而由(7.6.6)式可得

$$\boldsymbol{x}^{\mathrm{T}}[\nabla^2 f(\boldsymbol{x}_k)+\alpha_k\boldsymbol{I}_n]\boldsymbol{x}\geqslant 0; \tag{7.6.7}$$

若 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{d}_k=0$,令

$$\boldsymbol{x}_m=\boldsymbol{x}-\frac{1}{m}\boldsymbol{d}_k,m=1,2,\cdots,$$

则 $\lim\limits_{m\to\infty}\boldsymbol{x}_m=\boldsymbol{x}$,且由 $\boldsymbol{d}_k\neq\boldsymbol{0}$ 知 $\boldsymbol{x}_m^{\mathrm{T}}\boldsymbol{d}_k\neq 0$. 根据前面的讨论,有

$$\boldsymbol{x}_m^{\mathrm{T}}[\nabla^2 f(\boldsymbol{x}_k)+\alpha_k\boldsymbol{I}_n]\boldsymbol{x}_m\geqslant 0,m=1,2,\cdots,$$

在上式中,令 $m\to\infty$,即知(7.6.7)式也成立. 于是 $\nabla^2 f(\boldsymbol{x}_k)+\alpha_k\boldsymbol{I}_n$ 是半正定的.

充分性. 设 $\boldsymbol{d}_k$ 是问题(7.6.1)的可行点,且存在数 $\alpha_k\geqslant 0$,使得(7.6.4)式成立,而 $\nabla^2 f(\boldsymbol{x}_k)+\alpha_k\boldsymbol{I}_n$ 是半正定矩阵.

当 $\|\boldsymbol{d}_k\|<h_k$ 时,$\alpha_k=0$,故 $\nabla^2 f(\boldsymbol{x}_k)$ 是半正定矩阵,从而 $\varphi_k(\boldsymbol{y})$ 是凸函数,并且由(7.6.4)式有

$$\nabla\varphi_k(\boldsymbol{d}_k)=\nabla f(\boldsymbol{x}_k)+\nabla^2 f(\boldsymbol{x}_k)\boldsymbol{d}_k=\boldsymbol{0},$$

于是由定理3.1.4知,$\boldsymbol{d}_k$ 是凸函数 $\varphi_k(\boldsymbol{y})$ 的全局极小点,即知 $\boldsymbol{d}_k$ 为问题(7.6.1)的全局极小点.

当 $\|\boldsymbol{d}_k\|=h_k$ 时,根据(7.6.5)式,并注意到 $\nabla^2 f(\boldsymbol{x}_k)+\alpha_k\boldsymbol{I}_n$ 是半正定矩阵,$\alpha_k\geqslant 0$,因此,对于问题(7.6.1)的任何可行点 $\boldsymbol{y}$,即 $\|\boldsymbol{y}\|\leqslant h_k=\|\boldsymbol{d}_k\|$,均有 $\varphi_k(\boldsymbol{y})-\varphi_k(\boldsymbol{d}_k)\geqslant 0$. 这表明 $\boldsymbol{d}_k$ 是问题(7.6.1)的全局极小点. □

根据定理7.6.2,我们可以通过求解方程组

$$[\nabla^2 f(\boldsymbol{x}_k)+\alpha_k\boldsymbol{I}_n]\boldsymbol{y}=-\nabla f(\boldsymbol{x}_k)\ (\alpha_k\geqslant 0) \tag{7.6.8}$$

来确定问题(7.6.1)的解:如果矩阵 $\nabla^2 f(\boldsymbol{x}_k)+\alpha_k\boldsymbol{I}_n$ 正定,并且方程组(7.6.8)的解

$$\boldsymbol{d}_k=-[\nabla^2 f(\boldsymbol{x}_k)+\alpha_k\boldsymbol{I}_n]^{-1}\nabla f(\boldsymbol{x}_k)$$

满足 $\|\boldsymbol{d}_k\|\leqslant h_k$,那么 $\boldsymbol{d}_k$ 是问题(7.6.1)的全局极小点. 特别地,当 $\nabla^2 f(\boldsymbol{x}_k)$ 是正定矩阵,且 $\|[\nabla^2 f(\boldsymbol{x}_k)]^{-1}\nabla f(\boldsymbol{x}_k)\|\leqslant h_k$ 时,

$$\boldsymbol{d}_k=-[\nabla^2 f(\boldsymbol{x}_k)]^{-1}\nabla f(\boldsymbol{x}_k)$$

为凸规划问题(7.6.1)的惟一的全局极小点.

定理 7.6.3 设$\nabla^2 f(\boldsymbol{x}_k)+\alpha_k \boldsymbol{I}_n$是正定矩阵，则

$$\|\boldsymbol{d}_k\| = \|[\nabla^2 f(\boldsymbol{x}_k)+\alpha_k \boldsymbol{I}_n]^{-1}\nabla f(\boldsymbol{x}_k)\|$$

是α_k的单调减小函数.

证明 设对称矩阵$\nabla^2 f(x_k)$的特征值为$\mu_1,\mu_2,\cdots,\mu_n$，则存在正交矩阵$\boldsymbol{P}$，使得

$$\boldsymbol{P}^{\mathrm{T}} \nabla^2 f(\boldsymbol{x}_k)\boldsymbol{P}=\operatorname{diag}(\mu_1,\mu_2,\cdots,\mu_n).$$

记$\boldsymbol{P}^{\mathrm{T}}\nabla f(\boldsymbol{x}_k)=(z_1,z_2,\cdots,z_n)^{\mathrm{T}}$，类似于定理 7.5.3 中(2) 的证明，可得

$$\|\boldsymbol{d}_k\|^2=\sum_{i=1}^{n}\frac{z_i^2}{(\mu_i+\alpha_k)^2},$$

由此即知，$\|\boldsymbol{d}_k\|^2$从而$\|\boldsymbol{d}_k\|$是α_k的单调减小函数. □

由定理 7.6.3 可知，调整h_k可以通过调整α_k来实现：增大α_k相当于减小h_k；减小α_k相当于增大h_k.

根据信赖域法，我们给出信赖域 LM 法的具体步骤.

算法 7－11(信赖域 LM 法)

Step 1 选取初始数据. 选取初始点$\boldsymbol{x}_0$，选取$\alpha_0>0$，给定允许误差$\varepsilon>0$，$0<\beta<\gamma<1, 0<\theta<1, \rho>1$，令$k=0$.

Step 2 检查是否满足终止准则. 计算$\nabla f(\boldsymbol{x}_k)$和$\nabla^2 f(\boldsymbol{x}_k)$，若$\|\nabla f(\boldsymbol{x}_k)\|<\varepsilon$，迭代终止，$\boldsymbol{x}_k$为问题(7.1.1)的近似最优解.

Step 3 构造矩阵$\nabla^2 f(\boldsymbol{x}_k)+\alpha_k\boldsymbol{I}_n$，若该矩阵不正定，令$\alpha_k:=\rho\alpha_k$，重复Step 3；否则转 Step 4.

Step 4 求解方程组(7.6.8)，得出$\boldsymbol{d}_k$.

Step 5 计算$f(\boldsymbol{x}_k+\boldsymbol{d}_k)$和$r_k$的值.

Step 6 若$r_k<\beta$，则令$\alpha_{k+1}=\rho\alpha_k$；若$r_k>\gamma$，令$\alpha_{k+1}=\theta\alpha_k$；否则，令$\alpha_{k+1}=\alpha_k$.

Step 7 若$r_k\leqslant 0$，令$\boldsymbol{x}_{k+1}=\boldsymbol{x}_k$；否则，令$\boldsymbol{x}_{k+1}=\boldsymbol{x}_k+\boldsymbol{d}_k$.

Step 8 令$k:=k+1$，转 Step 2.

同样，信赖域 LM 法中的参数可按经验选取，一般取$\alpha_0=0.01,\beta=0.25,\gamma=0.75,\theta=0.5,\rho=4$.

习　题　七

1. 用最速下降法求解问题：

$$\min (x_1-2)^2+(x_1-2x_2)^2,$$

取初始点$\boldsymbol{x}_0=(0,3)^{\mathrm{T}}$，允许误差$\varepsilon=0.4$.

2. 用最速下降法求解正定二次函数的无约束最优化问题(7.1.5)，证明：

$$f(x_k)-f(x_{k+1})=\frac{[\nabla f(\boldsymbol{x}_k)^{\mathrm{T}}\nabla f(\boldsymbol{x}_k)]^2}{2\nabla f(\boldsymbol{x}_k)^{\mathrm{T}}\boldsymbol{Q}\nabla f(\boldsymbol{x}_k)},$$

其中 $\boldsymbol{x}_k$ 和 $\boldsymbol{x}_{k+1}$ 分别为第 k 次和第 $k+1$ 次迭代点.

3. 设 $f:\mathbb{R}^n\to\mathbb{R}$ 具有二阶连续偏导数,且对一切 $\boldsymbol{x}\in\mathbb{R}^n$,$\nabla^2 f(\boldsymbol{x})$ 为正定矩阵. 试说明把最速下降法应用于问题 $\min\limits_{\boldsymbol{x}\in\mathbb{R}^n} f(\boldsymbol{x})$ 时,得到近似迭代公式

$$\boldsymbol{x}_{k+1}=\boldsymbol{x}_k-\frac{\nabla f(\boldsymbol{x}_k)^{\mathrm{T}}\nabla f(\boldsymbol{x}_k)}{\nabla f(\boldsymbol{x}_k)^{\mathrm{T}}\nabla^2 f(\boldsymbol{x}_k)\nabla f(\boldsymbol{x}_k)}\nabla f(\boldsymbol{x}_k),k=0,1,2,\cdots.$$

试证这个公式给出的算法是下降算法. 并用给出的算法求解问题

$$\min\ 2x_1^2+(x_2-1)^4,$$

取初始点 $\boldsymbol{x}_0=(1,0)^{\mathrm{T}}$,迭代三次.

4. 用 Newton 法求解下列问题:

(1) $\min\ 2x_1^2+x_2^2+2x_1x_2+x_1-x_2$,

取初始点 $\boldsymbol{x}_0=(0,0)^{\mathrm{T}}$;

(2) $\min\ 2x_1^2+(x_2-1)^4$,

取初始点 $\boldsymbol{x}_0=(1,0)^{\mathrm{T}}$,允许误差 $\varepsilon=0.2$.

5. 用阻尼 Newton 法求解下列问题:

(1) $\min\dfrac{1}{(x_1^2+x_2^2+2)}$,

取初始点 $\boldsymbol{x}_0=(4,0)^{\mathrm{T}}$,允许误差 $\varepsilon=10^{-6}$;

(2) $\min(x_1-1)^4+(x_1-x_2)^2$,

取初始点 $\boldsymbol{x}_0=(0,0)^{\mathrm{T}}$,允许误差 $\varepsilon=10^{-6}$.

6. 试完成定理 7.2.2 的证明.

7. 设 $\boldsymbol{Q}\in\mathbb{R}^{n\times n}$ 为正定矩阵,$\boldsymbol{d}_1,\boldsymbol{d}_2,\cdots,\boldsymbol{d}_n\in\mathbb{R}^n$ 是一组线性无关向量,且

$$\boldsymbol{p}_k=\begin{cases}\boldsymbol{d}_k, & k=1,\\ \boldsymbol{d}_k-\sum\limits_{j=1}^{k-1}\dfrac{\boldsymbol{d}_j^{\mathrm{T}}\boldsymbol{Q}\boldsymbol{d}_k}{\boldsymbol{d}_j^{\mathrm{T}}\boldsymbol{Q}\boldsymbol{d}_j}\boldsymbol{d}_j, & k\geqslant 2.\end{cases}$$

试证:$\boldsymbol{p}_1,\boldsymbol{p}_2,\cdots,\boldsymbol{p}_n$ 是一组 $\boldsymbol{Q}$ 共轭向量.

8. 设 $\boldsymbol{Q}\in\mathbb{R}^{n\times n}$ 为正定矩阵,$\boldsymbol{d}_0,\boldsymbol{d}_1,\cdots,\boldsymbol{d}_{n-1}$ 是一组 $\boldsymbol{Q}$ 共轭的非零向量,证明:正定二次函数的无约束最优化问题(7.3.3)的最优解为

$$\bar{\boldsymbol{x}}=\sum_{i=0}^{n-1}\left(-\frac{\boldsymbol{d}_i^{\mathrm{T}}\boldsymbol{b}}{\boldsymbol{d}_i^{\mathrm{T}}\boldsymbol{Q}\boldsymbol{d}_i}\boldsymbol{d}_i\right).$$

9. 用 FR 共轭梯度法求解下列问题:

(1) $\min\ x_1^2+2x_2^2-2x_1x_2+2x_2+2$,

取初始点 $\boldsymbol{x}_0=(0,0)^{\mathrm{T}}$;

(2) $\min(x_1-1)^4+(x_1-x_2)^2$,

取初始点 $\boldsymbol{x}_0=(0,0)^{\mathrm{T}}$,允许误差 $\varepsilon=0.1$.

10. 用 DM 共轭梯度法和 PRP 共轭梯度法分别求解第 9 题中(2).

11. 试验证:DFP 公式和 BFGS 公式产生的迭代矩阵 $\boldsymbol{G}_{k+1}$ 满足拟 Newton 条件.

12. 设 $\boldsymbol{A}\in\mathbb{R}^{n\times n}$ 为正定矩阵,$\boldsymbol{u},\boldsymbol{v}\in\mathbb{R}^n$,且 $\boldsymbol{v}^{\mathrm{T}}\boldsymbol{A}^{-1}\boldsymbol{u}\neq -1$,试证 Sherman - Morrison 公式

$$(A+uv^{T})^{-1}=A^{-1}-\frac{A^{-1}uv^{T}A^{-1}}{1+v^{T}A^{-1}u}.$$

13. 用 DFP 法和 BFGS 法求解第 9 题.

14. 求解线性最小二乘问题

$$\min \|Ax-b\|^{2},$$

其中

$$A=\begin{bmatrix}1 & 1\\ 2 & -3\\ -1 & 4\end{bmatrix}, b=\begin{bmatrix}3\\ 2\\ 4\end{bmatrix}.$$

15. 设有非线性方程组

$$\begin{cases}x_1^2+2x_2^2-1=0,\\ 2x_1^2+x_2-2=0.\end{cases}$$

(1) 列出这个方程组的最小二乘问题的数学模型;

(2) 写出 Gauss - Newton 方向;

(3) 取初始点 $x_0=(2,2)^{T}$,用 Gauss - Newton 法迭代两次.

16. 用最小二乘 LM 法求解第 15 题中最小二乘问题,仍取初始点 $x_0=(2,2)^{T}$,并取初始参数 $\alpha_0=0.01$,放大因子 $\beta=10$,迭代两次.

17. 用信赖域 LM 法求解最优化问题:

$$\begin{cases}\min\ \varphi(y)=-2y_1+3y_2-4y_1^2-\frac{3}{2}y_2^2-6y_1y_2;\\ \text{s.t.}\ \ \|y\|\leqslant 1,\end{cases}$$

其中 $y=(y_1,y_2)^{T}$,取允许误差 $\varepsilon=0.001$.

第八章　无约束最优化的直接法

第七章介绍的无约束最优化的解析法,都要计算目标函数的一阶导数甚至二阶导数,但是,在有些无约束最优化问题中,目标函数的解析表达式比较复杂或者难以用明显的解析式表示出来,因而其导数很难求出或者无法求出,这就要求我们给出一些只涉及目标函数值的计算而不涉及目标函数的导数的求解方法. 这类仅仅利用目标函数值的信息直接建立搜索求解的方法通常称为直接搜索法(direct search method),简称为直接法. 尽管直接法因未利用目标函数导数的信息可能会影响算法的效果,但直接法构思直观,使用方便,效果稳定,实际工作者是很愿意采用的. 本章将介绍几种具有代表性的直接法.

8.1　坐标轮换法

坐标轮换法(univariate search technique)是 D'esopo(1959)提出来的,下面叙述它的基本原理.

考虑无约束最优化问题

$$\min f(\boldsymbol{x}), \tag{8.1.1}$$

其中 $\boldsymbol{x}\in\mathbb{R}^n, f: \mathbb{R}^n\to\mathbb{R}$.

给定 $\boldsymbol{x}_0\in\mathbb{R}^n$,先从点 $\boldsymbol{x}_0$ 出发沿第一个坐标轴方向 $\boldsymbol{e}_1=(1,0,\cdots,0)^{\mathrm{T}}$ 进行一维搜索,求 λ_0 和 $\boldsymbol{x}_1$,使得

$$f(\boldsymbol{x}_0+\lambda_0\boldsymbol{e}_1)=\min_{\lambda} f(\boldsymbol{x}_0+\lambda\boldsymbol{e}_1),$$

$$\boldsymbol{x}_1=\boldsymbol{x}_0+\lambda_0\boldsymbol{e}_1.$$

再从点 $\boldsymbol{x}_1$ 出发,沿第二个坐标轴方向 $\boldsymbol{e}_2=(0,1,0,\cdots,0)^{\mathrm{T}}$ 进行一维搜索,求 λ_1 和 $\boldsymbol{x}_2$,使得

$$f(\boldsymbol{x}_1+\lambda_1\boldsymbol{e}_2)=\min_{\lambda} f(\boldsymbol{x}_1+\lambda\boldsymbol{e}_2),$$

$$\boldsymbol{x}_2=\boldsymbol{x}_1+\lambda_1\boldsymbol{e}_2.$$

重复上述步骤,最后从点 $\boldsymbol{x}_{n-1}$ 出发,沿第 n 个坐标轴方向 $\boldsymbol{e}_n=(0,\cdots,0,1)^{\mathrm{T}}$ 进行一维搜索,求 λ_{n-1} 和 $\boldsymbol{x}_n$,使得

$$f(\boldsymbol{x}_{n-1}+\lambda_{n-1}\boldsymbol{e}_n)=\min_{\lambda} f(\boldsymbol{x}_{n-1}+\lambda\boldsymbol{e}_n),$$

$$x_n = x_{n-1} + \lambda_{n-1} e_n.$$

易知

$$f(x_j) \leqslant f(x_{j-1}), j = 1,2,\cdots,n. \tag{8.1.2}$$

设给定的允许误差 $\varepsilon > 0$,如果

$$\| x_n - x_0 \| < \varepsilon,$$

则迭代终止,x_n 为问题(8.1.1)的近似最优解;否则,重复上述步骤,即从点 x_n 出发,依次沿坐标轴方向 $e_1, e_2, \cdots, e_n$ 进行一维搜索.

需要指出的是,坐标轮换法中究竟应把哪个坐标轴方向放在前面、哪个坐标轴方向放在后面进行一维搜索,将会对算法的收敛速度产生较大的影响. 在实际问题中,可以按各因素(变量)对试验结果影响大小,依次自前往后排列坐标轴方向.

算法 8-1(坐标轮换法)

Step 1 选取初始数据. 选取初始点 x_0,给定允许误差 $\varepsilon > 0$,令 $k = 1$.

Step 2 进行一维搜索. 从点 x_{k-1} 出发,沿坐标轴方向 e_k 进行一维搜索,求 λ_{k-1} 和 x_k,使得

$$f(x_{k-1} + \lambda_{k-1} e_k) = \min_{\lambda} f(x_{k-1} + \lambda e_k),$$

$$x_k = x_{k-1} + \lambda_{k-1} e_k.$$

Step 3 检验迭代次数. 若 $k = n$,转 Step 4;否则,令 $k := k + 1$,返回 Step 2.

Step 4 检查是否满足终止准则. 若 $\| x_n - x_0 \| < \varepsilon$,终止迭代,得 x_n 为问题(8.1.1)的近似最优解;否则,令 $x_0 := x_n, k := 1$,返回 Step 2.

例 8.1.1 用坐标轮换法求解问题

$$\min f(x) = x_1^2 + x_2^2 - 3x_1 - x_1 x_2, \tag{8.1.3}$$

其中 $x = (x_1, x_2)^{\mathrm{T}}$. 取初始点 $x_0 = (0,0)^{\mathrm{T}}$,允许误差 $\varepsilon = 0.1$.

解 从点 x_0 出发沿 e_1 进行一维搜索,易得

$$\lambda_0 = \frac{3}{2}, x_1 = x_0 + \lambda_0 e_1 = \left(\frac{3}{2}, 0\right)^{\mathrm{T}};$$

从点 x_1 出发沿 e_2 进行一维搜索,得

$$\lambda_1 = \frac{3}{4}, x_2 = x_1 + \lambda_1 e_2 = \left(\frac{3}{2}, \frac{3}{4}\right)^{\mathrm{T}};$$

$$\| x_2 - x_0 \| > \varepsilon.$$

再从点 x_2 出发沿 e_1 进行一维搜索,得

$$\lambda_2 = \frac{3}{8}, x_3 = x_2 + \lambda_2 e_1 = \left(\frac{15}{8}, \frac{3}{4}\right)^{\mathrm{T}};$$

从点 x_3 出发沿 e_2 进行一维搜索,得

$$\lambda_3 = \frac{3}{16}, x_4 = x_3 + \lambda_3 e_2 = \left(\frac{15}{8}, \frac{15}{16}\right)^{\mathrm{T}};$$

$$\| x_4 - x_2 \| > \varepsilon.$$

再从点 x_4 出发沿 e_1 进行一维搜索,得

$$\lambda_4 = \frac{3}{32}, x_5 = x_4 + \lambda_4 e_1 = \left(\frac{63}{32}, \frac{15}{16}\right)^{\mathrm{T}};$$

从点 x_5 出发沿 e_2 进行一维搜索,得

$$\lambda_5 = \frac{3}{64}, x_6 = x_5 + \lambda_5 e_2 = \left(\frac{63}{32}, \frac{63}{64}\right)^{\mathrm{T}};$$

$$\| x_6 - x_4 \| > \varepsilon.$$

再从点 x_6 出发沿 e_1 进行一维搜索,得

$$\lambda_6 = \frac{3}{128}, x_7 = x_6 + \lambda_6 e_1 = \left(\frac{255}{128}, \frac{63}{64}\right)^{\mathrm{T}};$$

从点 x_7 出发沿 e_2 进行一维搜索,得

$$\lambda_7 = \frac{3}{256}, x_8 = x_7 + \lambda_7 e_2 = \left(\frac{255}{128}, \frac{255}{256}\right)^{\mathrm{T}};$$

$$\| x_8 - x_6 \| < \varepsilon.$$

迭代终止,得问题(8.1.3)的近似最优解为

$$x_8 = \left(\frac{255}{128}, \frac{255}{256}\right)^{\mathrm{T}}.$$

其实问题(8.1.3)的最优解为$(2,1)^{\mathrm{T}}$. □

下面给出坐标轮换法的收敛性定理,为此先证明一个引理.

引理 8.1.1 设 $f:\mathbb{R}^n \to \mathbb{R}$ 为连续函数. 若对于固定的 $j(1 \leqslant j \leqslant n)$,有

$$\min_{\lambda} f(y_k + \lambda e_j) = f(y_k + \lambda_k e_j), k = 1,2,\cdots, \tag{8.1.4}$$

记 $z_k(j) = y_k + \lambda_k e_j, k = 1,2,\cdots$,且有

$$\lim_{k\to\infty} y_k = \bar{y}, \lim_{k\to\infty} z_k(j) = \bar{z}_j,$$

则

$$\min_{\lambda} f(\bar{y} + \lambda e_j) = f(\bar{z}_j).$$

证明 由(8.1.4)式可知,对任意 $\lambda \in \mathbb{R}$,都有

$$f(z_k(j)) \leqslant f(y_k + \lambda e_j), k = 1,2,\cdots.$$

根据 f 的连续性,令 $k\to\infty$,对上式两边取极限,得

$$f(\bar{z}_j) \leqslant f(\bar{y} + \lambda e_j), \forall \lambda \in \mathbb{R}. \tag{8.1.5}$$

因为$\lim\limits_{k\to\infty} y_k = \bar{y}, \lim\limits_{k\to\infty} z_k(j) = \bar{z}_j$,且 $z_k(j) = y_k + \lambda_k e_j$,所以$\{\lambda_k\}$存在极限,设其极限为 $\bar{\lambda}$,则有 $\bar{z}_j = \bar{y} + \bar{\lambda} e_j$,从而由(8.1.5)式得:$\min\limits_{\lambda} f(\bar{y} + \lambda e_j) = f(\bar{z}_j)$. □

定理 8.1.2 设 $f:\mathbb{R}^n \to \mathbb{R}$ 具有一阶连续偏导数,$x_0 \in \mathbb{R}^n$,记 $\alpha = f(x_0)$,且水平集 $S(f,\alpha)$有界. 若$\{x_k\}$是用坐标轮换法求解问题(8.1.1)产生的点列,且在每次一维搜索中所得到的最优解都是惟一的,则

(1) 当$\{x_k\}$为有穷点列时,其最后一个点为f的平稳点;

(2) 当$\{x_k\}$为无穷点列时,它必有极限点,并且其任一极限点都是f的平稳点.

证明 (1) 当$\{x_k\}$为有穷点列时,由算法的终止条件可知,$\{x_k\}$的最后一个点x_m的下标必为n的正整数倍,且$\|x_m - x_{m-n}\| = 0$,即知$x_m = x_{m-n}$. 注意到点x_{m-n+j}是从点$x_{m-n+j-1}$出发沿e_j进行一维搜索得到的$(j=1,2,\cdots,n)$,从而由(8.1.2)式可得

$$f(x_m) \leqslant f(x_{m-1}) \leqslant \cdots \leqslant f(x_{m-n}),$$

于是,$f(x_m) = f(x_{m-j}), j=1,2,\cdots,n$. 根据定理中惟一性假设得知,$x_m = x_{m-1} = \cdots = x_{m-n}$,因此有

$$\min_\lambda f(x_m + \lambda e_j) = f(x_m), j = 1,2,\cdots,n.$$

由上式及定理6.1.1,知

$$\nabla f(x_m)^{\mathrm{T}} e_j = 0, j = 1,2,\cdots,n.$$

而$e_1,e_2,\cdots,e_n$是线性无关的,故由定理7.3.2,$\nabla f(x_m) = \mathbf{0}$.

(2) 当$\{x_k\}$为无穷点列时,由(8.1.2)式知$\{f(x_k)\}$为单调减少数列,且由

$$f(x_k) \leqslant f(x_0), k = 1,2,\cdots$$

可知,$\{x_k\} \subset S(f,\alpha)$. 又注意到$S(f,\alpha)$是有界闭集,故$\{x_k\}$为有界点列且连续函数$f$在$S(f,\alpha)$上有界,即知数列$\{f(x_k)\}$存在极限.

设$\bar{x}$为$\{x_k\}$的任一极限点,且设$\lim\limits_{k_m\to\infty} x_{k_m} = \bar{x}$,因为$\{x_{k_m}\}$是无穷点列,所以其中必有无穷多个点是沿某个方向$e_j$进行一维搜索得到的,不妨设$j=n$,$x_{k_m}$都是沿方向$e_n$进行一维搜索得到的. 现记$\bar{y}_1 = \bar{x}$,而且还有

$$\min_\lambda f(x_{k_m} + \lambda e_1) = f(x_{k_m} + \lambda_{k_m} e_1) = f(x_{k_m+1}).$$

由于$\{x_{k_m+1}\}$也是有界无穷点列,因此它必存在极限点,不妨设$x_{k_m+1} \to \bar{y}_2$,故$\lambda_{k_m} \to \bar{\lambda}_1$, $\bar{y}_2 = \bar{y}_1 + \bar{\lambda}_1 e_1$. 由引理8.1.1,可得

$$\min_\lambda f(\bar{y}_1 + \lambda e_1) = f(\bar{y}_2).$$

因f连续,且数列$\{f(x_k)\}$存在极限,故$f(\bar{y}_1) = f(\bar{y}_2)$.

由类似推理可知,对一切$j=1,2,\cdots,n$,都有

$$\min_\lambda f(\bar{y}_j + \lambda e_j) = f(\bar{y}_{j+1}),$$
$$f(\bar{y}_j) = f(\bar{y}_{j+1}),$$
$$\bar{y}_{j+1} = \bar{y}_j + \bar{\lambda}_j e_j.$$

由定理中惟一性假设可得

$$\bar{x} = \bar{y}_1 = \bar{y}_2 = \cdots = \bar{y}_n = \bar{y}_{n+1}.$$

于是有

$$\min_{\lambda} f(\bar{x} + \lambda e_j) = f(\bar{x}), j = 1,2,\cdots,n.$$

因此,根据(1)中相应的推理知,$\nabla f(\bar{x}) = \mathbf{0}$. □

8.2 模式搜索法

模式搜索法(pattern search method)是 Hooks 和 Jeeves(1961)提出来的. 这种方法的每一次迭代都是交替进行轴向移动(axis direction move)和模式移动(pattern move). 轴向移动的目的是探测下降的有利方向,而模式移动的目的则是沿着有利方向加速移动.

仍考虑无约束最优化问题(8.1.1),轴向移动是指从一个点 $y \in \mathbb{R}^n$ 出发,依次沿坐标轴方向 $e_1, e_2, \cdots, e_n$ 用定步长 δ_k 做探测性的搜索移动. 沿 $e_j(j=1,2,\cdots,n)$ 做探测性移动按以下格式进行:

(1) 正轴向探测 若 $f(y+\delta_k e_j) < f(y)$,探测成功,取 $y := y + \delta_k e_j$;否则,探测失败,再作负轴向探测;

(2) 负轴向探测 若 $f(y-\delta_k e_j) < f(y)$,探测成功,取 $y := y - \delta_k e_j$;否则,探测失败,y 不变.

将每一次探测性移动后得到的点,作为下一次探测性移动的开始点. 经过 n 次探测性移动,一般地可得到使 f 的值下降的点. 这就完成了一次轴向移动. 每次轴向移动的开始点称为参考点(consult point).

设 x_{k+1} 是以点 x_k 为参考点进行一次轴向移动所得到的点. 若

$$f(x_{k+1}) < f(x_k),$$

则从点 x_{k+1} 出发作模式移动. 否则,判断是否有 $\delta_k < \varepsilon$(其中 δ_k 为第 k 次轴向移动的步长,ε 为给定的允许误差)? 若有,迭代终止;若无,且 $x_{k+1} = x_k$,则缩短步长,仍从点 x_k 出发进行下一次轴向移动;若无,且 $x_{k+1} \neq x_k$,则从点 x_k 出发用步长 δ_k 进行下一次轴向移动.

从点 x_{k+1} 出发的模式移动是指以 1 为步长沿加速方向 $d_k = x_{k+1} - x_k$ 移动一步,得到新的参考点

$$y = x_{k+1} + d_k = 2x_{k+1} - x_k,$$

然后,从新的参考点 y 出发,仍以 δ_k 为步长进行轴向移动.

算法 8-2(模式搜索法)

Step 1 选取初始数据. 选取初始点 x_0,初始步长 $\delta_0 > 0$,给定收缩因子 $\alpha \in (0,1)$,给定允许误差 $\varepsilon > 0$,令 $k = 0$.

Step 2 确定参考点. 令 $y = x_k, j = 1$.

Step 3 进行正轴向探测. 从点 y 出发,沿 e_j 作正轴向探测:若 $f(y+\delta_k e_j) < f(y)$,令 $y := y + \delta_k e_j$,转 Step 5;否则,转 Step 4.

Step 4　进行负轴向探测. 从点 $\boldsymbol{y}$ 出发,沿 $\boldsymbol{e}_j$ 作负轴向探测:若 $f(\boldsymbol{y}-\delta_k\boldsymbol{e}_j)<f(\boldsymbol{y})$,令 $\boldsymbol{y}:=\boldsymbol{y}-\delta_k\boldsymbol{e}_j$,转 Step 5;否则,转 Step 5.

Step 5　检验探测次数. 若 $j<n$,令 $j:=j+1$,返回 Step 3;否则,令 $\boldsymbol{x}_{k+1}=\boldsymbol{y}$,转 Step 6.

Step 6　进行模式移动. 若 $f(\boldsymbol{x}_{k+1})<f(\boldsymbol{x}_k)$,从点 $\boldsymbol{x}_{k+1}$ 出发沿加速方向 $\boldsymbol{x}_{k+1}-\boldsymbol{x}_k$ 作模式移动,令

$$\boldsymbol{y}=2\boldsymbol{x}_{k+1}-\boldsymbol{x}_k,$$

$\delta_{k+1}=\delta_k$, $k:=k+1$, $j=1$,返回 Step 3;否则,转 Step 7.

Step 7　检查是否满足终止准则. 若 $\delta_k<\varepsilon$,迭代终止,得问题(8.1.1)的近似最优解为 $\boldsymbol{x}_k$;否则,转 Step 8.

Step 8　缩短步长. 若 $\boldsymbol{x}_{k+1}=\boldsymbol{x}_k$,令 $\delta_{k+1}=\alpha\delta_k$, $k:=k+1$,返回 Step 2;否则,令 $\boldsymbol{x}_{k+1}=\boldsymbol{x}_k$, $\delta_{k+1}=\delta_k$, $k:=k+1$,返回 Step 2.

例 8.2.1　用模式搜索法求解问题(8.1.3),取初始点 $\boldsymbol{x}_0=(0,0)^{\mathrm{T}}$,初始步长 $\delta_0=1$,收缩因子 $\alpha=0.25$,允许误差 $\varepsilon=0.1$.

解　第一次迭代:

令初始参考点 $\boldsymbol{y}=\boldsymbol{x}_0=(0,0)^{\mathrm{T}}$,因为 $f(\boldsymbol{y})=0$, 有

$$f(\boldsymbol{y}+\delta_0\boldsymbol{e}_1)=-2<f(\boldsymbol{y}),$$

所以取 $\boldsymbol{y}:=\boldsymbol{y}+\delta_0\boldsymbol{e}_1=(1,0)^{\mathrm{T}}$. 又有

$$f(\boldsymbol{y}+\delta_0\boldsymbol{e}_2)=-2=f(\boldsymbol{y}),$$
$$f(\boldsymbol{y}-\delta_0\boldsymbol{e}_2)=0>f(\boldsymbol{y}),$$

故第一次轴向移动结束,令 $\boldsymbol{x}_1=\boldsymbol{y}=(1,0)^{\mathrm{T}}$. 由于 $f(\boldsymbol{x}_1)<f(\boldsymbol{x}_0)$,因此进行模式移动,得下一个参考点

$$\boldsymbol{y}=2\boldsymbol{x}_1-\boldsymbol{x}_0=(2,0)^{\mathrm{T}},$$

取 $\delta_1=\delta_0$.

第二次迭代:

因为 $f(\boldsymbol{y})=-2$,有

$$f(\boldsymbol{y}+\delta_1\boldsymbol{e}_1)=0>f(\boldsymbol{y}),$$
$$f(\boldsymbol{y}-\delta_1\boldsymbol{e}_1)=-2=f(\boldsymbol{y}),$$
$$f(\boldsymbol{y}+\delta_1\boldsymbol{e}_2)=-3<f(\boldsymbol{y}),$$

所以取 $\boldsymbol{y}:=\boldsymbol{y}+\delta_1\boldsymbol{e}_2=(2,1)^{\mathrm{T}}$. 第二次轴向移动结束,令 $\boldsymbol{x}_2=\boldsymbol{y}=(2,1)^{\mathrm{T}}$. 由于 $f(\boldsymbol{x}_2)<f(\boldsymbol{x}_1)$,因此进行模式移动,得下一参考点

$$\boldsymbol{y}=2\boldsymbol{x}_2-\boldsymbol{x}_1=(3,2)^{\mathrm{T}},$$

取 $\delta_2=\delta_1$.

第三次迭代:

因为$f(\boldsymbol{y}) = -2$,有

$$f(\boldsymbol{y}+\delta_2\boldsymbol{e}_1) = 0 > f(\boldsymbol{y}),$$
$$f(\boldsymbol{y}-\delta_2\boldsymbol{e}_1) = -2 = f(\boldsymbol{y}),$$
$$f(\boldsymbol{y}+\delta_2\boldsymbol{e}_2) = 0 > f(\boldsymbol{y}),$$
$$f(\boldsymbol{y}-\delta_2\boldsymbol{e}_2) = -2 = f(\boldsymbol{y}),$$

所以第三次轴向移动结束,令 $\boldsymbol{x}_3=\boldsymbol{y}=(3,2)^{\mathrm{T}}$. 由于$f(\boldsymbol{x}_3)>f(\boldsymbol{x}_2)$,$\delta_2=1>0.1$,且 $\boldsymbol{x}_3\neq\boldsymbol{x}_2$,因此令 $\boldsymbol{x}_3=\boldsymbol{x}_2=(2,1)^{\mathrm{T}}$,$\delta_3=\delta_2$,取参考点 $\boldsymbol{y}=\boldsymbol{x}_3=(2,1)^{\mathrm{T}}$.

第四次迭代:

因为$f(\boldsymbol{y}) = -3$,有

$$f(\boldsymbol{y}+\delta_3\boldsymbol{e}_1) = -2 > f(\boldsymbol{y}),$$
$$f(\boldsymbol{y}-\delta_3\boldsymbol{e}_1) = -2 > f(\boldsymbol{y}),$$
$$f(\boldsymbol{y}+\delta_3\boldsymbol{e}_2) = -2 > f(\boldsymbol{y}),$$
$$f(\boldsymbol{y}-\delta_3\boldsymbol{e}_2) = -2 > f(\boldsymbol{y}),$$

所以第四次轴向移动结束,令 $\boldsymbol{x}_4=\boldsymbol{y}=(2,1)^{\mathrm{T}}$. 注意到$f(\boldsymbol{x}_4)=f(\boldsymbol{x}_3)$,从而要检查步长大小. 因 $\delta_3=1>0.1$,且 $\boldsymbol{x}_4=\boldsymbol{x}_3$,故缩短步长,令 $\delta_4=\alpha\delta_3=0.25$,取参考点 $\boldsymbol{y}=\boldsymbol{x}_4=(2,1)^{\mathrm{T}}$.

第五次迭代:

与第四次迭代类似有 $\boldsymbol{x}_5=\boldsymbol{x}_4$,再缩短步长,令 $\delta_5=\alpha\delta_4=0.062\ 5$,取参考点 $\boldsymbol{y}=\boldsymbol{x}_5=(2,1)^{\mathrm{T}}$.

第六次迭代:

与第五次迭代类似有 $\boldsymbol{x}_6=\boldsymbol{x}_5$,但此时 $\delta_5=0.062\ 5<0.1$,迭代终止,得问题(8.1.3)的近似最优解 $\boldsymbol{x}_6=(2,1)^{\mathrm{T}}$. 其实 $\boldsymbol{x}_6$ 就是问题(8.1.3)的最优解. □

虽然模式搜索法可能要计算很多点的函数值才能得到目标函数的近似极小点,但是它易于在计算机上实现,实际效果也不错,因此被认为是一种可靠的方法.

下面给出模式搜索法的收敛性定理.

设$\{\boldsymbol{x}_k\}$和$\{\delta_k\}$分别为模式搜索法求解问题(8.1.1)产生的迭代点列和步长数列,则由 δ_k 的取法知,$\{\delta_k\}$是单调减小的且以 0 为下界,故存在极限. 从而根据算法的终止条件知

$$\lim_{k\to\infty}\delta_k = 0. \tag{8.2.1}$$

又由算法的迭代过程中 $\boldsymbol{x}_k$ 的定义,有

$$f(\boldsymbol{x}_{k+1}) \leqslant f(\boldsymbol{x}_k),k = 0,1,2,\cdots. \tag{8.2.2}$$

定理 8.2.1 设$f:\mathbb{R}^n\to\mathbb{R}$ 是具有一阶连续偏导数的凸函数,$\boldsymbol{x}_0\in\mathbb{R}^n$,记 $\alpha=f(\boldsymbol{x}_0)$,并且水平集 $S(f,\alpha)$有界. 若$\{\boldsymbol{x}_k\}$为由模式搜索法求解问题(8.1.1)产生的点列,则$\{\boldsymbol{x}_k\}$必存在极限点,且其任一极限点都是问题(8.1.1)的最优解.

证明　先证点列$\{\boldsymbol{x}_k\}$至少存在一个极限点为问题(8.1.1)的最优解.

由(8.2.1)式易知,$\{\boldsymbol{x}_k\}$中必有无穷多个k_i,使

$$\boldsymbol{x}_{k_i}=\boldsymbol{x}_{k_i+1}=\cdots=\boldsymbol{x}_{k_i+n},\tag{8.2.3}$$

其中$\boldsymbol{x}_{k_i+j}$是从点$\boldsymbol{x}_{k_i+j-1}$出发沿$\boldsymbol{e}_j$进行探测性移动后得到的点$(j=1,2,\cdots,n)$.

由(8.2.2)式知,$\{f(\boldsymbol{x}_k)\}$为单调减小数列,从而$\{\boldsymbol{x}_k\}\subseteq S(f,\alpha)$,所以由$S(f,\alpha)$的有界性知$\{\boldsymbol{x}_k\}$有界.

设$\{\boldsymbol{x}_{k_i'}\}$是有界点列$\{\boldsymbol{x}_{k_i}\}$的收敛子列,且

$$\lim_{k_i'\to\infty}\boldsymbol{x}_{k_i'}=\bar{\boldsymbol{x}}.$$

又由(8.2.1)式可知,对于任给的$\bar{\delta}>0$,存在正整数$\bar{k}$,使

$$0<\delta_k<\bar{\delta},k\geqslant\bar{k}.\tag{8.2.4}$$

由$\{\boldsymbol{x}_{k_i'}\}$的取法及(8.2.3)式,有

$$f(\boldsymbol{x}_{k_i'+j-1}\pm\delta_{k_i'+j-1}\boldsymbol{e}_j)\geqslant f(\boldsymbol{x}_{k_i'+j-1}),j=1,2,\cdots,n.\tag{8.2.5}$$

从而,当$k_i'\geqslant\bar{k}$时,有

$$f(\boldsymbol{x}_{k_i'+j-1}\pm\bar{\delta}\boldsymbol{e}_j)\geqslant f(\boldsymbol{x}_{k_i'+j-1}),j=1,2,\cdots n.\tag{8.2.6}$$

若不然,设有某个$j(1\leqslant j\leqslant n)$,使

$$f(\boldsymbol{x}_{k_i'+j-1}\pm\bar{\delta}\boldsymbol{e}_j)<f(\boldsymbol{x}_{k_i'+j-1}).$$

根据(8.2.4)式可知,存在$\lambda\in(0,1)$,使得

$$\delta_{k_i'+j-1}=\lambda\bar{\delta},k'_i\geqslant\bar{k},$$

于是,由f的凸性,有

$$\begin{aligned}f(x_{k_i'+j-1}\pm\delta_{k_i'+j-1}\boldsymbol{e}_j)&=f((1-\lambda)\boldsymbol{x}_{k_i'+j-1}+\lambda(\boldsymbol{x}_{k_i'+j-1}\pm\bar{\delta}\boldsymbol{e}_j))\\&\leqslant(1-\lambda)f(\boldsymbol{x}_{k_i'+j-1})+\lambda f(\boldsymbol{x}_{k_i'+j-1}\pm\bar{\delta}\boldsymbol{e}_j)<f(\boldsymbol{x}_{k_i'+j-1}),\end{aligned}$$

此与(8.2.5)式矛盾.

由(8.2.3)式有$\boldsymbol{x}_{k_i'}=\boldsymbol{x}_{k_i'+j},j=1,2,\cdots,n$. 因此由(8.2.6)式即得

$$f(\boldsymbol{x}_{k_i'}\pm\bar{\delta}\boldsymbol{e}_j)\geqslant f(\boldsymbol{x}_{k_i'}),j=1,2,\cdots,n.$$

从而,由f的连续性和上式可知,对一切$\bar{\delta}>0$,有

$$f(\bar{\boldsymbol{x}}\pm\bar{\delta}\boldsymbol{e}_j)\geqslant f(\bar{\boldsymbol{x}}),j=1,2,\cdots,n.$$

这表示

$$\min_{\lambda}f(\bar{\boldsymbol{x}}+\lambda\boldsymbol{e}_j)=f(\bar{\boldsymbol{x}}),j=1,2,\cdots,n.$$

由此知

$$\nabla f(\bar{\boldsymbol{x}})^{\mathrm{T}}\boldsymbol{e}_j=0,j=1,2,\cdots,n.$$

而$\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_n$线性无关,故由定理7.3.2知,$\nabla f(\bar{\boldsymbol{x}})=\boldsymbol{0}$. 根据推论3.1.5,$\bar{\boldsymbol{x}}$为问题(8.1.1)的最优解.

再证$\{\boldsymbol{x}_k\}$的任一极限点$\tilde{\boldsymbol{x}}$都是问题(8.1.1)的最优解.

设$\lim\limits_{k_i\to\infty}\boldsymbol{x}_{k_i}=\tilde{\boldsymbol{x}}$,由上知,$\{f(\boldsymbol{x}_k)\}$为单调减小数列,且以$f(\bar{\boldsymbol{x}})$为下界,故

$\{f(\boldsymbol{x}_k)\}$存在极限. 从而由 f 的连续性有

$$\lim_{k\to\infty} f(\boldsymbol{x}_k) = \lim_{k_i\to\infty} f(\boldsymbol{x}_{k_i}) = f(\tilde{\boldsymbol{x}}) = f(\bar{\boldsymbol{x}}),$$

所以,$\tilde{\boldsymbol{x}}$ 也是问题(8.1.1)的最优解. □

8.3 旋转方向法

旋转方向法(rotating direction method)是 Rosenbrock(1960)提出来的. 这种方法是在每一次迭代中采用变步长的轴向移动. 然后利用轴向的旋转来产生一组新的方向作为下一次迭代的轴向,其目的是为了加快收敛速度.

在模式搜索法中,轴向移动的步长在一定阶段是保持不变的,只有当满足某种条件时才缩短步长. 在旋转方向法中,其轴向移动的步长在探测过程中独立地随时变化,即按照探测的成功要增大步长,又随探测的失败而缩短步长. 具体做法如下:预先给定一个步长收缩因子 $\alpha\in(0,1)$ 和一个步长放大因子 $\beta>1$,在从点 $\boldsymbol{x}_k$ 到点 $\boldsymbol{x}_{k+1}$ 的轴向移动中,先从 $\boldsymbol{x}_k$ 作为参考点 $\boldsymbol{y}$ 出发,依次沿作为轴向的 n 个单位正交向量 $\boldsymbol{d}_1,\boldsymbol{d}_2,\cdots,\boldsymbol{d}_n$ 做变步长 $\delta_1^{(k)},\delta_2^{(k)},\cdots,\delta_n^{(k)}$ 的探测性移动. 对于一切 $j=1,2,\cdots,n$,若 $f(\boldsymbol{y}+\delta_j^{(k)}\boldsymbol{d}_j)<f(\boldsymbol{y})$,则探测成功,令

$$\boldsymbol{y} := \boldsymbol{y} + \delta_j^{(k)}\boldsymbol{d}_j,\ \delta_j^{(k)} := \beta\delta_j^{(k)};$$

若 $f(\boldsymbol{y}+\delta_j^{(k)}\boldsymbol{d}_j)\geqslant f(\boldsymbol{y})$,则探测失败,令

$$\delta_j^{(k)} := -\alpha\delta_j^{(k)}.$$

在完成一次轴向移动之后,把本次轴向移动最后的那个点 $\boldsymbol{y}$ 作为参考点开始新的一次轴向移动,如此继续下去,直到某一次轴向移动中沿 n 个轴向的探测均告失败,此时第 $k+1$ 次迭代的探测阶段结束. 把探测阶段结束时的那个 $\boldsymbol{y}$ 点记为 $\boldsymbol{x}_{k+1}$,它也是下一次迭代探测阶段的第一个参考点.

为了提高求解效率,需要进行各轴向的旋转,并以新的 n 个单位正交方向组作为下一次迭代的轴向. 容易想到,从探测阶段的开始点指向该探测阶段的结束点的方向 $\boldsymbol{x}_{k+1}-\boldsymbol{x}_k$ 很可能是有利于目标函数值下降的方向,因此新的一组轴向应该包含这个方向.

根据探测阶段结果,有

$$\boldsymbol{x}_{k+1} - \boldsymbol{x}_k = \lambda_1\boldsymbol{d}_1 + \lambda_2\boldsymbol{d}_2 + \cdots + \lambda_n\boldsymbol{d}_n,$$

其中 λ_j 是探测阶段中所有沿方向 $\boldsymbol{d}_j$ 移动的步长的代数和. 对一切 $j=1,2,\cdots,n$,令

$$\boldsymbol{p}_j = \begin{cases} \boldsymbol{d}_j, & \lambda_j = 0, \\ \sum_{i=j}^{n}\lambda_i\boldsymbol{d}_i, & \lambda_j \neq 0. \end{cases} \tag{8.3.1}$$

再利用 Gram - Schmidt 正交化方法将方向组 $\boldsymbol{p}_1,\boldsymbol{p}_2,\cdots,\boldsymbol{p}_n$ 进行标准正交化. 对一切 $j=1,2,\cdots,n$,作

$$\boldsymbol{q}_j=\begin{cases}\boldsymbol{p}_j, & j=1,\\ \boldsymbol{p}_j-\sum_{i=1}^{j-1}\dfrac{\boldsymbol{q}_i^{\mathrm{T}}\boldsymbol{p}_j}{\boldsymbol{q}_i^{\mathrm{T}}\boldsymbol{q}_i}\boldsymbol{q}_i, & j\geqslant 2,\end{cases} \tag{8.3.2}$$

然后单位化,令

$$\bar{\boldsymbol{d}}_j=\frac{\boldsymbol{q}_j}{\|\boldsymbol{q}_j\|},j=1,2,\cdots,n. \tag{8.3.3}$$

所得标准正交化的方向 $\bar{\boldsymbol{d}}_1,\bar{\boldsymbol{d}}_2,\cdots,\bar{\boldsymbol{d}}_n$ 可作为第 $k+2$ 次迭代的一组轴向.

定理 8.3.1 假设 $\boldsymbol{d}_1,\boldsymbol{d}_2,\cdots,\boldsymbol{d}_n$ 是一组单位正交向量,则对任何一组 $\lambda_1,\lambda_2,\cdots,\lambda_n$,由(8.3.1) ~ (8.3.3)式定义的方向 $\bar{\boldsymbol{d}}_1,\bar{\boldsymbol{d}}_2,\cdots,\bar{\boldsymbol{d}}_n$ 也是单位正交向量.

证明 因 $\boldsymbol{d}_1,\boldsymbol{d}_2,\cdots,\boldsymbol{d}_n$ 是一组正交的单位向量,故由定理 7.3.1 知,$\boldsymbol{d}_1,\boldsymbol{d}_2,\cdots,\boldsymbol{d}_n$ 是线性无关的. 假定存在一组数 $\mu_1,\mu_2,\cdots,\mu_n$,使 $\sum_{j=1}^{n}\mu_j\boldsymbol{p}_j=\boldsymbol{0}$,记

$$I=\{j\mid\lambda_j=0\},J(j)=\{i\mid i\notin I,i\leqslant j\},$$

则由(8.3.1)式有

$$\begin{aligned}\boldsymbol{0}=\sum_{j=1}^{n}\mu_j\boldsymbol{p}_j&=\sum_{j\in I}\mu_j\boldsymbol{d}_j+\sum_{j\notin I}\mu_j\Big(\sum_{i=j}^{n}\lambda_i\boldsymbol{d}_i\Big)\\&=\sum_{j\in I}\mu_j\boldsymbol{d}_j+\sum_{j\notin I}\Big(\lambda_j\sum_{i\in J(j)}\mu_i\Big)\boldsymbol{d}_j.\end{aligned}$$

由于 $\boldsymbol{d}_1,\boldsymbol{d}_2,\cdots,\boldsymbol{d}_n$ 是线性无关的,因此

$$\mu_j=0(j\in I),\lambda_j\sum_{i\in J(j)}\mu_i=0(j\notin I).$$

但当 $j\notin I$ 时,$\lambda_j\neq0$,从而 $\sum_{i\in J(j)}\mu_i=0$,故由 $J(j)$ 的定义知 $\mu_j=0$,所以 $\mu_1=\mu_2=\cdots=\mu_n=0$. 这就证明了 $\boldsymbol{p}_1,\boldsymbol{p}_2,\cdots,\boldsymbol{p}_n$ 线性无关.

根据 Gram - Schmidt 正交化方法可知,由(8.3.2)和(8.3.3)式产生的 $\bar{\boldsymbol{d}}_1,\bar{\boldsymbol{d}}_2,\cdots,\bar{\boldsymbol{d}}_n$ 是一组正交的单位向量. □

算法 8 - 3(旋转方向法)

Step 1 选取初始数据. 选取初始点 $\boldsymbol{x}_0$,初始单位正交方向组 $\boldsymbol{d}_1,\boldsymbol{d}_2,\cdots,\boldsymbol{d}_n$(可取 $\boldsymbol{d}_1,\boldsymbol{d}_2,\cdots,\boldsymbol{d}_n$ 为坐标轴方向 $\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots\boldsymbol{e}_n$). 给定初始步长 $\boldsymbol{\delta}_0=(\delta_1^{(0)},\delta_2^{(0)},\cdots,\delta_n^{(0)})^{\mathrm{T}}>\boldsymbol{0}$,收缩因子 $\alpha\in(0,1)$,放大因子 $\beta>1$,允许误差 $\varepsilon>0$,令 $k=0$.

Step 2 确定参考点. 取参考点 $\boldsymbol{y}=x_k$,并令 $\boldsymbol{z}=\boldsymbol{x}_k,j=1$.

Step 3 进行轴向探测. 若 $f(\boldsymbol{y}+\delta_j^{(k)}\boldsymbol{d}_j)<f(\boldsymbol{y})$,令

$$\boldsymbol{y}:=\boldsymbol{y}+\delta_j^{(k)}\boldsymbol{d}_j,\delta_j^{(k)}:=\beta\,\delta_j^{(k)}$$

转 Step 4;否则,令 $\delta_j^{(k)}:=-\alpha\delta_j^{(k)}$,转 Step 4.

Step 4　检验探测次数. 若 $j < n$, 令 $j := j+1$, 返回 Step 3; 否则, 转 Step 5.

Step 5　判断探测是否结束. 若 $f(\boldsymbol{y}) < f(\boldsymbol{z})$, 令 $\boldsymbol{z} = \boldsymbol{y}, j = 1$, 返回 Step 3; 若

$$f(\boldsymbol{y}) = f(\boldsymbol{z}), f(\boldsymbol{y}) < f(\boldsymbol{x}_k),$$

令 $\boldsymbol{x}_{k+1} = \boldsymbol{y}$, 转 Step 6; 若

$$f(\boldsymbol{y}) = f(\boldsymbol{z}) = f(\boldsymbol{x}_k),$$

转 Step 7.

Step 6　检查是否满足终止准则. 若 $\| \boldsymbol{x}_{k+1} - \boldsymbol{x}_k \| < \varepsilon$, 迭代终止, $\boldsymbol{x}_{k+1}$ 为问题(8.1.1)的近似最优解; 否则, 转 Step 8.

Step 7　检验步长大小. 若对一切 $j = 1, 2, \cdots, n$, $| \delta_j^{(k)} | < \varepsilon$, 迭代终止, $\boldsymbol{x}_k$ 为问题(8.1.1)的近似最优解; 否则, 令 $\boldsymbol{z} = \boldsymbol{y}, j = 1$, 返回 Step 3.

Step 8　进行轴向旋转. 计算 $\lambda_1, \lambda_2, \cdots, \lambda_n$, 利用(8.3.1) ~ (8.3.3)式构造新的单位正交方向 $\bar{\boldsymbol{d}}_1, \bar{\boldsymbol{d}}_2, \cdots, \bar{\boldsymbol{d}}_n$, 并令

$$\boldsymbol{d}_j = \bar{\boldsymbol{d}}_j, \delta_j^{(k+1)} = \delta_j^{(k)}, j = 1, 2, \cdots, n,$$

$k := k+1$, 返回 Step 2.

例 8.3.1　用旋转方向法求解问题(8.1.3), 取初始点 $\boldsymbol{x}_0 = (0,0)^{\mathrm{T}}$, 初始搜索方向 $\boldsymbol{d}_1 = (1,0)^{\mathrm{T}}, \boldsymbol{d}_2 = (0,1)^{\mathrm{T}}$, 初始步长 $\boldsymbol{\delta}_0 = (1,1)^{\mathrm{T}}$, 收缩因子 $\alpha = 0.5$, 放大因子 $\beta = 3$, 试求 $\boldsymbol{x}_2$ 和转轴后的轴向.

解　第一次迭代:

令初始参考点 $\boldsymbol{y} = \boldsymbol{x}_0, \boldsymbol{z} = \boldsymbol{x}_0$. 因为 $f(\boldsymbol{y}) = 0$, 有

$$f(\boldsymbol{y} + \delta_1^{(0)} \boldsymbol{d}_1) = -2 < f(\boldsymbol{y}),$$

所以, 令 $\boldsymbol{y} := \boldsymbol{y} + \delta_1^{(0)} \boldsymbol{d}_1 = (1,0)^{\mathrm{T}}, \delta_1^{(0)} := \beta\, \delta_1^{(0)} = 3$. 又有

$$f(\boldsymbol{y} + \delta_2^{(0)} \boldsymbol{d}_2) = -2 = f(\boldsymbol{y}),$$

故令 $\delta_2^{(0)} := -\alpha \delta_2^{(0)} = -0.5$. 由于 $f(\boldsymbol{y}) < f(\boldsymbol{z})$, 因此令 $\boldsymbol{z} = \boldsymbol{y}$, 继续作轴向移动. 因

$$f(\boldsymbol{y} + \delta_1^{(0)} \boldsymbol{d}_1) = 4 > f(\boldsymbol{y}),$$

故令 $\delta_1^{(0)} := -\alpha \delta_1^{(0)} = -1.5$. 又有

$$f(\boldsymbol{y} + \delta_2^{(0)} \boldsymbol{d}_2) = -1.25 > f(\boldsymbol{y}),$$

故令 $\delta_2^{(0)} := -\alpha \delta_2^{(0)} = 0.25$. 由于 $f(\boldsymbol{y}) = f(\boldsymbol{z}), f(\boldsymbol{y}) < f(\boldsymbol{x}_0)$, 因此探测阶段结束, 令 $\boldsymbol{x}_1 = \boldsymbol{y} = (1,0)^{\mathrm{T}}$.

因 $\| \boldsymbol{x}_1 - \boldsymbol{x}_0 \| > \varepsilon$, 再作轴向旋转. 先求各轴向移动的步长代数和, 求得 $\lambda_1 = 1, \lambda_2 = 0$, 按(8.3.1)式有

$$\boldsymbol{p}_2 = \boldsymbol{d}_2, \boldsymbol{p}_1 = \lambda_1 \boldsymbol{d}_1 + \lambda_2 \boldsymbol{d}_2 = \boldsymbol{d}_1.$$

即下次迭代的轴向仍为 $\boldsymbol{d}_1 = (1,0)^{\mathrm{T}}, \boldsymbol{d}_2 = (0,1)^{\mathrm{T}}$, 并取 $\boldsymbol{\delta}_1 = \boldsymbol{\delta}_0 = (-1.5, 0.25)^{\mathrm{T}}$.

第二次迭代：

以 $y=(1,0)^T$ 为参考点，令 $z=(1,0)^T$，重新开始轴向移动. 因为

$$f(y+\delta_1^{(1)}d_1)=1.75>f(y),$$

所以取 $\delta_1^{(1)}:=-\alpha\delta_1^{(1)}=0.75$. 又有

$$f(y+\delta_2^{(1)}d_2)=-2.1875<f(y),$$

故令 $y:=y+\delta_2^{(1)}d_2=(1,0.25)^T,\delta_2^{(1)}:=\beta\,\delta_2^{(1)}=0.75$. 由于 $f(y)<f(z)$，因此令 $z=y$. 继续作轴向移动. 因为

$$f(y+\delta_1^{(1)}d_1)=-2.5625<f(y),$$

故令 $y:=y+\delta_1^{(1)}d_1=(1.75,0.25)^T,\delta_1^{(1)}:=\beta\,\delta_1^{(1)}=2.25$. 又有

$$f(y+\delta_2^{(1)}d_2)=-2.9375<f(y),$$

故令 $y:=y+\delta_2^{(1)}d_1=(1.75,1)^T,\delta_2^{(1)}=\beta\,\delta_2^{(1)}=2.25$. 由于 $f(y)<f(z)$，因此令 $z=y$. 继续作轴向移动. 因为

$$f(y+\delta_1^{(1)}d_1)=1>f(y),$$

所以取 $\delta_1^{(1)}:=-\alpha\delta_1^{(1)}=-1.125$. 又有

$$f(y+\delta_2^{(1)}d_2)=1.5625>f(y),$$

故令 $\delta_2^{(1)}:=-\alpha\delta_2^{(1)}=-1.125$. 由于 $f(y)=f(z),f(y)<f(x_1)$，因此探测阶段结束，令 $x_2=y=(1.75,1)^T$.

因 $\|x_2-x_1\|>\varepsilon$，再进行轴向旋转. 先求各轴向移动的步长的代数和，得 $\lambda_1=0.75,\lambda_2=1$，按(8.3.1)式有

$$p_1=\lambda_1d_1+\lambda_2d_2=(0.75,1)^T,$$

$$p_2=\lambda_2d_2=(0,1)^T.$$

从而由(8.3.2)和(8.3.3)式得到新的一组轴向

$$d_1=(0.6,0.8)^T,d_2=(-0.8,0.6)^T$$ □

由于旋转方向法的转轴条件和转轴后所取的步长不甚合理，因此方法的收敛性问题一直没有解决. Davies, Swann 和 Campey 将旋转方向法中沿轴向的变步长移动改为沿轴向做最优一维搜索，给出了修正的旋转方向法，而修正的旋转方向法是收敛的（参见文献[16,48]）.

8.4 Powell 法

本节介绍由 Powell(1964)提出的一种求解无约束最优化问题(8.1.1)的直接法，它本质上是以正定二次函数为背景，以共轭方向为基础的一种方法.

8.4.1 原始 Powell 法

考虑正定二次函数的无约束最优化问题

$$\min f(\boldsymbol{x}) = \frac{1}{2}\boldsymbol{x}^{\mathrm{T}}\boldsymbol{Q}\boldsymbol{x} + \boldsymbol{b}^{\mathrm{T}}\boldsymbol{x} + c, \tag{8.4.1}$$

其中 $\boldsymbol{x} \in \mathbb{R}^n, \boldsymbol{Q} \in \mathbb{R}^{n\times n}$ 为正定矩阵，$\boldsymbol{b} \in \mathbb{R}^n, c \in \mathbb{R}$.

定理 8.4.1 若分别从两个点 $\boldsymbol{y}_0, \tilde{\boldsymbol{y}}_0 \in \mathbb{R}^n$ 出发，依次沿 k 个非零的 $\boldsymbol{Q}$ 共轭方向 $\boldsymbol{d}_0, \boldsymbol{d}_1, \cdots, \boldsymbol{d}_{k-1}$ 对问题(8.4.1)做最优一维搜索得到点 $\boldsymbol{y}_k, \tilde{\boldsymbol{y}}_k \in \mathbb{R}^n$，则向量 $\boldsymbol{y}_k - \tilde{\boldsymbol{y}}_k$ 与 $\boldsymbol{d}_j (j=0,1,\cdots k-1)$ 是 $\boldsymbol{Q}$ 共轭的.

证明 因为 $\boldsymbol{y}_k$ 和 $\tilde{\boldsymbol{y}}_k$ 是分别从 $\boldsymbol{y}_0$ 点和 $\tilde{\boldsymbol{y}}_0$ 出发，依次沿非零的 $\boldsymbol{Q}$ 共轭方向 $\boldsymbol{d}_0, \boldsymbol{d}_1, \cdots, \boldsymbol{d}_{k-1}$ 做一维搜索得到的点，所以由定理 7.3.3 的证明中的(7.3.8)式有

$$\nabla f(\boldsymbol{y}_k)^{\mathrm{T}}\boldsymbol{d}_j = 0, j = 0,1,\cdots,k-1,$$
$$\nabla f(\tilde{\boldsymbol{y}}_k)^{\mathrm{T}}\boldsymbol{d}_j = 0, j = 0,1,\cdots,k-1,$$

即

$$(\boldsymbol{Q}\boldsymbol{y}_k + \boldsymbol{b})^{\mathrm{T}}\boldsymbol{d}_j = 0, j = 0,1,\cdots,k-1,$$
$$(\boldsymbol{Q}\tilde{\boldsymbol{y}}_k + \boldsymbol{b})^{\mathrm{T}}\boldsymbol{d}_j = 0, j = 0,1,\cdots,k-1.$$

上面两式相减，得

$$(\boldsymbol{y}_k - \tilde{\boldsymbol{y}}_k)^{\mathrm{T}}\boldsymbol{Q}\boldsymbol{d}_j = 0, j = 0,1,\cdots,k-1.$$

这就证明了 $\boldsymbol{y}_k - \tilde{\boldsymbol{y}}_k$ 与 $\boldsymbol{d}_j (j=0,1,\cdots,k-1)$ 是 $\boldsymbol{Q}$ 共轭的. □

根据定理 8.4.1，Powell 给出了如下的一个算法，称之为原始 Powell 法. 它是由所谓基本搜索、加速搜索和调整搜索方向三个部分组成.

算法 8-4(原始 Powell 法)

Step 1 选取初始数据. 选取初始点 $\boldsymbol{x}_0$，n 个线性无关的初始搜索方向 $\boldsymbol{d}_0, \boldsymbol{d}_1, \cdots, \boldsymbol{d}_{n-1}$，给定允许误差 $\varepsilon > 0$，令 $k = 0$.

Step 2 进行基本搜索. 令 $\boldsymbol{y}_0 = \boldsymbol{x}_k$，依次沿 $\boldsymbol{d}_0, \boldsymbol{d}_1, \cdots \boldsymbol{d}_{n-1}$ 进行一维搜索. 对一切 $j = 1,2,\cdots,n$，记

$$f(\boldsymbol{y}_{j-1} + \lambda_{j-1}\boldsymbol{d}_{j-1}) = \min_{\lambda} f(\boldsymbol{y}_{j-1} + \lambda\boldsymbol{d}_{j-1}),$$
$$\boldsymbol{y}_j = \boldsymbol{y}_{j-1} + \lambda_{j-1}\boldsymbol{d}_{j-1}.$$

Step 3 进行加速搜索. 取加速方向 $\boldsymbol{d}_n = \boldsymbol{y}_n - \boldsymbol{y}_0$，若 $\|\boldsymbol{d}_n\| < \varepsilon$，迭代终止，得 $\boldsymbol{y}_n$ 为问题的近似最优解；否则，从点 $\boldsymbol{y}_n$ 出发沿 $\boldsymbol{d}_n$ 进行一维搜索，求出 λ_n，使得

$$f(\boldsymbol{y}_n + \lambda_n\boldsymbol{d}_n) = \min_{\lambda} f(\boldsymbol{y}_n + \lambda\boldsymbol{d}_n).$$

记 $\boldsymbol{x}_{k+1} = \boldsymbol{y}_n + \lambda_n\boldsymbol{d}_n$，转 Step 4.

Step 4 调整搜索方向. 在原来 n 个方向 $\boldsymbol{d}_0, \boldsymbol{d}_1, \cdots, \boldsymbol{d}_{n-1}$ 中，除去 $\boldsymbol{d}_0$ 增添 $\boldsymbol{d}_n$，构成新的搜索方向，即令

$$\boldsymbol{d}_j := \boldsymbol{d}_{j+1}, j = 0, 1, \cdots, n-1.$$

令 $k := k+1$，返回 Step 2.

为了加快收敛速度，原始 Powell 法在沿 n 个线性无关的方向作一维搜索之后，增加了沿加速方向的一维搜索，并且利用加速方向去调整下一次迭代的搜索方向.

定理 8.4.2 设用原始 Powell 法求解问题(8.4.1). 若迭代已进行了 $m(\leqslant n)$ 次，且每一次迭代后为下一次迭代确定的前 n 个搜索方向都是线性无关的，则 m 次迭代产生的 m 个加速方向是一组 $\boldsymbol{Q}$ 共轭的非零方向.

证明 设 n 个线性无关的初始搜索方向为 $\boldsymbol{d}_0, \boldsymbol{d}_1, \cdots, \boldsymbol{d}_{n-1}$，$m$ 次迭代中相继产生的加速方向依次为

$$\boldsymbol{d}_n, \boldsymbol{d}_{n+1}, \cdots, \boldsymbol{d}_{n+m-1}. \tag{8.4.2}$$

现在对 m 用归纳法证明(8.4.2)式是一组 $\boldsymbol{Q}$ 共轭的非零方向.

当 $m=2$ 时，设第二次迭代中从点 $\boldsymbol{y}_0$ 出发做基本搜索产生的点依次为 $\boldsymbol{y}_1, \boldsymbol{y}_2, \cdots, \boldsymbol{y}_n$，从而得加速方向为

$$\boldsymbol{d}_{n+1} = \boldsymbol{y}_n - \boldsymbol{y}_0.$$

因为 $\boldsymbol{y}_0$ 是第一次迭代中最后沿加速方向 $\boldsymbol{d}_n$ 做一维搜索得到的点，$\boldsymbol{y}_n$ 是第二次迭代中从 $\boldsymbol{y}_{n-1}$ 出发沿方向 $\boldsymbol{d}_n$ 作一维搜索得到的点，所以由定理 8.4.1 知，$\boldsymbol{d}_{n+1}$ 与 $\boldsymbol{d}_n$ 是 $\boldsymbol{Q}$ 共轭的. 又由于第二次迭代为下一次迭代确定的搜索方向 $\boldsymbol{d}_2, \cdots, \boldsymbol{d}_{n-1}, \boldsymbol{d}_n, \boldsymbol{d}_{n+1}$ 是线性无关的，因此 $\boldsymbol{d}_n$ 与 $\boldsymbol{d}_{n+1}$ 是两个非零的 $\boldsymbol{Q}$ 共轭方向.

设 $m=r$ 时定理成立，即

$$\boldsymbol{d}_n, \boldsymbol{d}_{n+1}, \cdots, \boldsymbol{d}_{n+r-1}$$

是一组 $\boldsymbol{Q}$ 共轭的非零方向. 注意到 $\boldsymbol{d}_{n+r}$ 的起点和终点都是依次沿上述 r 个 $\boldsymbol{Q}$ 共轭方向做一维搜索得到的点，于是由定理 8.4.1 知

$$\boldsymbol{d}_n, \boldsymbol{d}_{n+1}, \cdots, \boldsymbol{d}_{n+r-1}, \boldsymbol{d}_{n+r} \tag{8.4.3}$$

是 Q 共轭的方向组. 又因第 $r+1$ 次迭代为下一次迭代确定的搜索方向组

$$\boldsymbol{d}_{r+1}, \boldsymbol{d}_{r+2}, \cdots, \boldsymbol{d}_{n+r-1}, \boldsymbol{d}_{n+r}$$

线性无关，故 $\boldsymbol{d}_{n+r} \neq \boldsymbol{0}$，即(8.4.3)式是一组非零的 $\boldsymbol{Q}$ 共轭方向. 这表明 $m=r+1$ 时定理成立. □

从定理 8.4.2 可以看出，原始 Powell 法也是一种共轭方向法. 由于它仅仅需要计算目标函数值而不必求其导数值，因此，原始 Powell 法比 7.3 节中共轭方向法(包括共轭梯度法)更具实用性. 另一方面，若将原始 Powell 法用于求解问题(8.4.1)，则当迭代 n 次得到点 $\boldsymbol{x}_n$，且每次迭代的前 n 个搜索方向都线性无关时，由于 $\boldsymbol{x}_n$ 是依次沿 n 个 $\boldsymbol{Q}$ 共轭的非零方向

$$\boldsymbol{d}_n, \boldsymbol{d}_{n+1}, \cdots, \boldsymbol{d}_{2n-1}$$

做一维搜索得到的点，因此根据定理 7.3.3，$\boldsymbol{x}_n$ 必是问题(8.4.1)的最优解. 这表

明,如果每次迭代的前 n 个搜索方向都线性无关,则原始 Powell 法具有二次终止性.

虽然原始 Powell 法是以求解问题(8.4.1)为背景建立起来的,但算法中并未涉及目标函数 f 的 Hesse 矩阵 $\boldsymbol{Q}$,所以原始 Powell 法也可以用来求解一般无约束最优化问题(8.1.1).

例 8.4.1 用原始 Powell 法求解问题(8.1.3),取初始点 $\boldsymbol{x}_0=(0,0)^{\mathrm{T}}$,初始搜索方向组 $\boldsymbol{d}_0=(0,1)^{\mathrm{T}},\boldsymbol{d}_1=(1,0)^{\mathrm{T}}$.

解 第一次迭代:

令 $\boldsymbol{y}_0=\boldsymbol{x}_0=(0,0)^{\mathrm{T}}$,从点 $\boldsymbol{y}_0$ 出发沿 $\boldsymbol{d}_0$ 进行一维搜索,易得

$$\lambda_0=0,\boldsymbol{y}_1=\boldsymbol{y}_0+\lambda_0\boldsymbol{d}_0=(0,0)^{\mathrm{T}}.$$

接着从 $\boldsymbol{y}_1$ 出发沿 $\boldsymbol{d}_1$ 进行一维搜索,得

$$\lambda_1=\frac{3}{2},\boldsymbol{y}_2=\boldsymbol{y}_1+\lambda_1\boldsymbol{d}_1=\left(\frac{3}{2},0\right)^{\mathrm{T}}.$$

再从点 $\boldsymbol{y}_2$ 出发沿加速方向 $\boldsymbol{d}_2=\boldsymbol{y}_2-\boldsymbol{y}_0=\left(\frac{3}{2},0\right)^{\mathrm{T}}$ 做一维搜索,得

$$\lambda_2=0,\boldsymbol{x}_1=\boldsymbol{y}_2+\lambda_2\boldsymbol{d}_2=\left(\frac{3}{2},0\right)^{\mathrm{T}}.$$

第二次迭代:

两个基本搜索方向为

$$\boldsymbol{d}_0=(1,0)^{\mathrm{T}},\boldsymbol{d}_1=\left(\frac{3}{2},0\right)^{\mathrm{T}}.$$

注意到这两个搜索方向的第二个分量均为 0,因此沿这两个方向进行一维搜索得到的点的第二个分量将保持为 0,这样永远达不到问题的最优解 $\bar{\boldsymbol{x}}=(2,1)^{\mathrm{T}}$,其原因是 $\boldsymbol{d}_0,\boldsymbol{d}_1$ 线性相关. 这说明运用原始 Powell 法求解本例失败. □

由此可见,在原始 Powell 法中,保持每次迭代中前 n 个搜索方向线性无关是十分重要的. 而且,Powell 本人已经注意到,原始 Powell 法即使不像例 8.4.1 那样极端的情况,但也可能产生近乎线性相关的搜索方向,特别是变量很多时更是如此.

8.4.2 Powell 法

为了避免出现搜索方向组线性相关的现象,Powell 及其他人对原始 Powell 法进行了修正. 他们首先研究了搜索方向组 $\boldsymbol{d}_0,\boldsymbol{d}_1,\cdots,\boldsymbol{d}_{n-1}$ 的 $\boldsymbol{Q}$ 共轭性度量.

定理 8.4.3 设 $\boldsymbol{Q}\in\mathbb{R}^{n\times n}$ 为正定矩阵. 若向量组 $\boldsymbol{d}_0,\boldsymbol{d}_1,\cdots,\boldsymbol{d}_{n-1}$ 满足

$$\boldsymbol{d}_j^{\mathrm{T}}\boldsymbol{Q}\boldsymbol{d}_j=1,j=0,1,\cdots,n-1, \tag{8.4.4}$$

则 $\boldsymbol{d}_0,\boldsymbol{d}_1,\cdots,\boldsymbol{d}_{n-1}$ 是一组 $\boldsymbol{Q}$ 共轭的向量的充要条件为 $|\det\boldsymbol{D}|$ 取得最大值,其中

矩阵 $D=(d_0,d_1,\cdots,d_{n-1})$.

证明 必要性. 设方向 $p_0,p_1,\cdots,p_{n-1}$ 满足:对于一切 $i,j=0,1\cdots,n-1$,有

$$p_i^{\mathrm{T}}Qp_j=\begin{cases}0, i\neq j,\\1, i=j.\end{cases}\tag{8.4.5}$$

记 $P=(p_0,p_1,\cdots,p_{n-1})$,我们只需证明:对于任何满足(8.4.4)式的向量组 d_0,$d_1,\cdots,d_{n-1}$,有

$$|\det D|\leqslant|\det P|.\tag{8.4.6}$$

根据定理7.3.1,Q 共轭的非零向量组 $p_0,p_1,\cdots,p_{n-1}$ 是线线无关的,因此它们是$\mathbb{R}^n$ 的一组基,于是,我们可将每个 d_j 用 $p_0,p_1,\cdots,p_{n-1}$ 的线性组合表示为 $d_j=\sum\limits_{i=0}^{n-1}u_{ij}p_i$. 记

$$U=\begin{bmatrix}u_{00} & u_{01} & \cdots & u_{0,n-1}\\u_{10} & u_{11} & \cdots & u_{1,n-1}\\\vdots & \vdots & & \vdots\\u_{n-1,0} & u_{n-1,1} & \cdots & u_{n-1,n-1}\end{bmatrix}=(u_0,u_1,\cdots,u_{n-1}),$$

则 $d_j=Pu_j(j=0,1,\cdots,n-1)$,即知 $D=PU$,从而有

$$|\det D|=|\det P|\cdot|\det U|.\tag{8.4.7}$$

由(8.4.5)式知 $P^{\mathrm{T}}QP=I_n$,故对一切 $j=0,1,\cdots,n-1$,有

$$d_j^{\mathrm{T}}Qd_j=(Pu_j)^{\mathrm{T}}Q(Pu_j)=u_j^{\mathrm{T}}P^{\mathrm{T}}QPu_j=u_j^{\mathrm{T}}u_j,$$

于是,由(8.4.4)式有

$$u_j^{\mathrm{T}}u_j=1, j=0,1,\cdots,n-1,\tag{8.4.8}$$

因为 $U^{\mathrm{T}}U$ 是半正定矩阵,所以由 Hadamard 不等式得

$$0\leqslant\det(U^{\mathrm{T}}U)\leqslant\prod_{j=0}^{n-1}u_j^{\mathrm{T}}u_j.$$

故由(8.4.8)式知,$|\det U|\leqslant1$,代入(8.4.7)式得(8.4.6)式.

充分性. 设(8.4.4)式成立,而且 $|\det D|$ 取得最大值,用反证法证明 d_0,$d_1,\cdots,d_{n-1}$ 是 Q 共轭的. 若不然,不妨设 d_0,d_1 不是 Q 共轭的,则 $d_0^{\mathrm{T}}Qd_1=c\neq0$. 因为 Q 是正定矩阵,所以存在满秩矩阵 L,使 $Q=L^{\mathrm{T}}L$,从而由 Cauchy 不等式和(8.4.4)式得

$$\begin{aligned}|c|&=|(Ld_0)^{\mathrm{T}}(Ld_1)|\\&\leqslant\|Ld_0\|\cdot\|Ld_1\|\\&=((Ld_0)^{\mathrm{T}}(Ld_0))^{1/2}((Ld_1)^{\mathrm{T}}(Ld_1))^{1/2}\\&=(d_0^{\mathrm{T}}Qd_0)^{1/2}(d_1^{\mathrm{T}}Qd_1)^{1/2}=1,\end{aligned}$$

并且 $|c|=1$ 的充要条件是 d_0 与 d_1 线性相关,即$\det D=0$,此时,$|\det D|$ 没有取得最大值. 下设 $0<|c|<1$. 定义

$$\tilde{d}_0 = \frac{d_0 - c d_1}{(1-c^2)^{1/2}},$$

$$\widetilde{D} = (\tilde{d}_0, \tilde{d}_1, \cdots, \tilde{d}_{n-1}),$$

容易验证 $\tilde{d}_0, d_1, d_2, \cdots, d_{n-1}$,满足(8.4.4)式,并且

$$|\det\widetilde{D}| = \frac{1}{(1-c^2)^{1/2}} |\det D| > |\det D|,$$

即知 $|\det D|$ 未取得最大值,与条件矛盾. □

根据定理8.4.3,若方向组 $d_0, d_1, \cdots, d_{n-1}$ 满足(8.4.4),则可用 $|\det D|$ 作为方向组 $d_0, d_1, \cdots, d_{n-1}$ 关于 Q 的共轭性度量:$|\det D|$ 越大,方向组的 Q 共轭性就越好.

利用上述 Q 共轭性度量修改原始 Powell 法的原则是使每一次迭代中 n 个搜索方向的 Q 共轭性度量不减少. 此时,需要解决下面两个问题:

(1) 每次迭代中搜索方向 $d_0, d_1, \cdots, d_{n-1}$ 的 Q 共轭性度量如何给出?

(2) 在确定下一次迭代的 n 个搜索方向时,以加速方向 d_n 替代 $d_0, d_1, \cdots, d_{n-1}$ 中哪一个才能使 Q 共轭性度量最大?

下面我们针对问题(8.4.1)来回答上述两个问题.

对于第一个问题,设本次迭代中 n 个搜索方向为 $d_0, d_1, \cdots, d_{n-1}$,让每一个方向在 Q 度量意义下单位化,即令

$$d_j := \frac{d_j}{\sqrt{d_j^{\mathrm{T}} Q d_j}}, j = 0, 1 \cdots, n-1.$$

记 $D = (d_0, d_1, \cdots, d_{n-1})$,则 $|\det D|$ 就是本次迭代中前 n 个搜索方向的 Q 共轭性度量.

对于第二个问题,设本次迭代中从 y_0 点出发,依次沿方向 $d_0, d_1, \cdots, d_{n-1}$ 做一维搜索得到点 $y_1, y_2, \cdots, y_n$,则

$$\begin{aligned} y_n &= y_{n-1} + \lambda_{n-1} d_{n-1} = y_{n-2} + \lambda_{n-2} d_{n-2} + \lambda_{n-1} d_{n-1} \\ &= \cdots = y_0 + \lambda_0 d_0 + \lambda_1 d_1 + \cdots + \lambda_{n-1} d_{n-1}, \end{aligned}$$

即加速方向 $d_n = y_n - y_0$ 可表示为

$$d_n = \lambda_0 d_0 + \lambda_1 d_1 + \cdots + \lambda_{n-1} d_{n-1}, \tag{8.4.9}$$

其中 $d_0, d_1, \cdots, d_{n-1}$ 满足

$$d_j^{\mathrm{T}} Q d_j = 1, j = 0, 1, \cdots, n-1. \tag{8.4.10}$$

再把 d_n 在 Q 度量意义下单位化,即令

$$\begin{aligned} \beta &= \sqrt{d_n^{\mathrm{T}} Q d_n} = \sqrt{(y_n - y_0)^{\mathrm{T}} Q (y_n - y_0)}, \\ d_n &:= d_n / \beta, \end{aligned} \tag{8.4.11}$$

并记 $\alpha_j = \lambda_j / \beta, j = 0, 1, \cdots, n-1$,则由(8.4.9)式有

$$d_n = \alpha_0 d_0 + \alpha_1 d_1 + \cdots + \alpha_{n-1} d_{n-1}. \tag{8.4.12}$$

若以加速方向 d_n 替换某个 $d_m(0 \leqslant m \leqslant n-1)$，记

$$\widetilde{D} = (d_0, \cdots, d_{m-1}, d_n, d_{m+1}, \cdots, d_{n-1}), \tag{8.4.13}$$

则下一次迭代的 n 个搜索方向

$$d_0, \cdots, d_{m-1}, d_{m+1}, \cdots, d_{n-1}, d_n$$

的 Q 共轭性度量为 $|\det\widetilde{D}|$. 把(8.4.12)式代入(8.4.13)式，由行列式运算性质可以得到：$|\det\widetilde{D}| = |\alpha_m| \cdot |\det D|$. 由此可知，要使 $|\det\widetilde{D}|$ 尽可能大，m 的选取应使

$$|\alpha_m| = \max_{0 \leqslant j \leqslant n-1} |\alpha_j|. \tag{8.4.14}$$

要使下一次迭代中 n 个搜索方向的 Q 共轭性度量增大，必有

$$|\alpha_m| > 1. \tag{8.4.15}$$

因为对一切 $j = 1, 2, \cdots, n$，有

$$y_j - y_{j-1} = \lambda_{j-1} d_{j-1} = \beta\alpha_{j-1} d_{j-1},$$

所以，利用正定二次函数 f 的 Taylor 展开得

$$f(y_{j-1}) = f(y_j) + \nabla f(y_j)^{\mathrm{T}}(y_{j-1} - y_j) + \frac{1}{2}(y_{j-1} - y_j)^{\mathrm{T}} Q (y_{j-1} - y_j)$$

$$= f(y_j) - \beta\alpha_{j-1} \nabla f(y_j)^{\mathrm{T}} d_{j-1} + \frac{1}{2}(\beta\alpha_{j-1})^2 d_{j-1}^{\mathrm{T}} Q d_{j-1}.$$

注意到最优一维搜索导致 $\nabla f(y_j)^{\mathrm{T}} d_{j-1} = 0$，从而由(8.4.10)式有

$$(\beta\alpha_{j-1})^2 = 2[f(y_{j-1}) - f(y_j)], j = 1, 2, \cdots, n. \tag{8.4.16}$$

因此，以 d_n 替代 d_m 的条件(8.4.14)等价于

$$f(y_m) - f(y_{m+1}) = \max_{0 \leqslant j \leqslant n-1} \{f(y_j) - f(y_{j+1})\} \tag{8.4.17}$$

上式表明，被 d_n 替换的 d_m 应是 n 个方向 $d_0, d_1, \cdots, d_{n-1}$ 中使目标函数值下降最多的那个方向.

利用(8.4.16)式可知，以 d_n 替换 d_m 的另一个条件(8.4.15)等价于

$$\beta^2 < 2[f(y_m) - f(y_{m+1})],$$

而由正定二次函数 f 的 Taylor 展开有

$$\begin{aligned}
\beta^2 &= (y_n - y_0)^{\mathrm{T}} Q (y_n - y_0) \\
&= \left[\nabla f(y_n)^{\mathrm{T}}(y_0 - y_n) + \frac{1}{2}(y_0 - y_n)^{\mathrm{T}} Q (y_0 - y_n)\right] + \\
&\quad \left[\nabla f(y_n)^{\mathrm{T}}(y_n - y_0) + \frac{1}{2}(y_n - y_0)^{\mathrm{T}} Q (y_n - y_0)\right] \\
&= [f(y_0) - f(y_n)] + [f(2y_n - y_0) - f(y_n)] \\
&= f(y_0) - 2f(y_n) + f(2y_n - y_0),
\end{aligned}$$

于是,条件(8.4.15)又等价于

$$f(\boldsymbol{y}_0)-2f(\boldsymbol{y}_n)+f(2\boldsymbol{y}_n-\boldsymbol{y}_0)<2[f(\boldsymbol{y}_m)-f(\boldsymbol{y}_{m+1})]. \qquad (8.4.18)$$

由上面的分析可知,加速方向 $\boldsymbol{d}_n$ 是否要替换原来的 n 个搜索方向之一,以及替换哪一个方向,可由(8.4.17)和(8.4.18)两式决定. 对原始 Powell 法做这样的修改得到的方法就称为 Powell 法.

算法 8-5(Powell 法)

Step 1　选取初始数据. 选取初始点 $\boldsymbol{x}_0$,n 个线性无关的初始搜索方向 $\boldsymbol{d}_0$, $\boldsymbol{d}_1,\cdots,\boldsymbol{d}_{n-1}$,给定允许误差 $\varepsilon>0$,令 $k=0$.

Step 2　进行基本搜索. 令 $\boldsymbol{y}_0=\boldsymbol{x}_k$,依次沿 $\boldsymbol{d}_0,\boldsymbol{d}_1,\cdots,\boldsymbol{d}_{n-1}$ 进行一维搜索. 对一切 $j=1,2,\cdots,n$,记

$$f(\boldsymbol{y}_{j-1}+\lambda_{j-1}\boldsymbol{d}_{j-1})=\min_{\lambda} f(\boldsymbol{y}_{j-1}+\lambda\boldsymbol{d}_{j-1}),$$

$$\boldsymbol{y}_j=\boldsymbol{y}_{j-1}+\lambda_{j-1}\boldsymbol{d}_{j-1}.$$

Step 3　检查是否满足终止准则. 取加速方向 $\boldsymbol{d}_n=\boldsymbol{y}_n-\boldsymbol{y}_0$,若 $\|\boldsymbol{d}_n\|<\varepsilon$,迭代终止,得 $\boldsymbol{y}_n$ 为问题的近似最优解;否则,转 Step 4.

Step 4　确定搜索方向. 按(8.4.17)式确定 m,若(8.4.18)式成立,转 Step 5;否则,转 Step 6.

Step 5　调整搜索方向. 从点 $\boldsymbol{y}_n$ 出发沿方向 $\boldsymbol{d}_n$ 进行一维搜索,求出 λ_n,使得

$$f(\boldsymbol{y}_n+\lambda_n\boldsymbol{d}_n)=\min_{\lambda} f(\boldsymbol{y}_n+\lambda\boldsymbol{d}_n)$$

令 $\boldsymbol{x}_{k+1}=\boldsymbol{y}_n+\lambda_n\boldsymbol{d}_n$,再令

$$\boldsymbol{d}_j:=\boldsymbol{d}_{j+1},j=m,m+1,\cdots,n-1.$$

$k:=k+1$,返回 Step 2.

Step 6　不调整搜索方向. 令 $\boldsymbol{x}_{k+1}=\boldsymbol{y}_n$,$k:=k+1$,返回 Step 2.

需要说明以下三点:

(1) 在 Powell 法的计算步骤中,搜索方向组 $\boldsymbol{d}_0,\boldsymbol{d}_1,\cdots,\boldsymbol{d}_{n-1}$ 及加速方向 $\boldsymbol{d}_n$ 都没有在 $\boldsymbol{Q}$ 度量意义下单位化,但这并不影响计算,因为这些单位化因子

$$\sqrt{\boldsymbol{d}_j^{\mathrm{T}}\boldsymbol{Q}\boldsymbol{d}_j},j=0,1,\cdots,n$$

可以包含在一维搜索的步长中.

(2) 与原始 Powell 法一样,Powell 法也可以用来求解一般无约束最优化问题(8.1.1).

(3) 由于初始搜索方向 $\boldsymbol{d}_0,\boldsymbol{d}_1,\cdots,\boldsymbol{d}_{n-1}$ 是线性无关的,且迭代中搜索方向组的 $\boldsymbol{Q}$ 共轭性度量不会减少,因此在 Powell 法中每一次迭代的前 n 个搜索方向都是线性无关的. 但是,Powell 法不再具有二次终止性.

例 8.4.2　用 Powell 法求解问题(8.1.3),仍取初始点 $\boldsymbol{x}_0=(0,0)^{\mathrm{T}}$,初始搜索方向组

$$\boldsymbol{d}_0=(0,1)^{\mathrm{T}},\boldsymbol{d}_1=(1,0)^{\mathrm{T}}.$$

给定允许误差 $\varepsilon=0.1$.

解 第一次迭代:

同例 8.4.1 一样,得

$$\boldsymbol{y}_1=(0,0)^{\mathrm{T}},\boldsymbol{y}_2=\left(\frac{3}{2},0\right)^{\mathrm{T}},$$

$$\boldsymbol{d}_2=\boldsymbol{y}_2-\boldsymbol{y}_0=\left(\frac{3}{2},0\right)^{\mathrm{T}}.$$

因为 $\|\boldsymbol{d}_2\|=3/2>\varepsilon$,所以要确定调整方向. 由于

$$f(\boldsymbol{y}_0)=0,f(\boldsymbol{y}_1)=0,f(\boldsymbol{y}_2)=-\frac{9}{4},$$

按(8.4.17)式有

$$f(\boldsymbol{y}_1)-f(\boldsymbol{y}_2)=\max\{f(\boldsymbol{y}_j)-f(\boldsymbol{y}_{j+1})\mid j=0,1\},$$

因此 $m=1$,并且

$$f(\boldsymbol{y}_m)-f(\boldsymbol{y}_{m+1})=f(\boldsymbol{y}_1)-f(\boldsymbol{y}_2)=\frac{9}{4}.$$

又因 $f(2\boldsymbol{y}_2-\boldsymbol{y}_0)=0$,故(8.4.18)式不成立. 于是,不调整搜索方向组,并令 $\boldsymbol{x}_1=\boldsymbol{y}_2=\left(\frac{3}{2},0\right)^{\mathrm{T}}$.

第二次迭代:

取 $\boldsymbol{y}_0=\boldsymbol{x}_1=\left(\frac{3}{2},0\right)^{\mathrm{T}}$,从点 $\boldsymbol{y}_0$ 出发沿 $\boldsymbol{d}_0$ 做一维搜索,得

$$\lambda_0=\frac{3}{4},\boldsymbol{y}_1=\boldsymbol{y}_0+\lambda_0\boldsymbol{d}_0=\left(\frac{3}{2},\frac{3}{4}\right)^{\mathrm{T}}.$$

接着从点 $\boldsymbol{y}_1$ 出发沿 $\boldsymbol{d}_1$ 作一维搜索,得

$$\lambda_1=\frac{3}{8},\boldsymbol{y}_2=\boldsymbol{y}_1+\lambda_1\boldsymbol{d}_1=\left(\frac{15}{8},\frac{3}{4}\right)^{\mathrm{T}}.$$

由此有加速方向

$$\boldsymbol{d}_2=\boldsymbol{y}_2-\boldsymbol{y}_0=\left(\frac{3}{8},\frac{3}{4}\right)^{\mathrm{T}}.$$

因为 $\|\boldsymbol{d}_2\|=3\sqrt{5}/8>\varepsilon$,所以要确定调整方向. 因

$$f(\boldsymbol{y}_0)=-\frac{9}{4},f(\boldsymbol{y}_1)=-\frac{45}{16},f(\boldsymbol{y}_2)=-\frac{189}{64},$$

故按(8.4.17)式易知 $m=0$,并且

$$f(\boldsymbol{y}_m)-f(\boldsymbol{y}_{m+1})=f(\boldsymbol{y}_0)-f(\boldsymbol{y}_1)=\frac{9}{16}.$$

由于

$$f(2\boldsymbol{y}_2-\boldsymbol{y}_0)=-\frac{45}{16},$$

因此(8.4.18)式成立. 于是,从点 $\boldsymbol{y}_2$ 出发沿 $\boldsymbol{d}_2$ 做一维搜索,得

$$\lambda_2=\frac{1}{3},\boldsymbol{x}_2=\boldsymbol{y}_2+\lambda_2\boldsymbol{d}_2=(2,1)^{\mathrm{T}}.$$

同时,以 $\boldsymbol{d}_2$ 替换 $\boldsymbol{d}_0$,即下一次迭代的搜索方向组取为

$$\boldsymbol{d}_0=(1,0)^{\mathrm{T}},\boldsymbol{d}_1=\left(\frac{3}{8},\frac{3}{4}\right)^{\mathrm{T}}.$$

第三次迭代:

取 $\boldsymbol{y}_0=\boldsymbol{x}_2=(2,1)^{\mathrm{T}}$,类似地可求得

$$\boldsymbol{y}_1=(2,1)^{\mathrm{T}},\boldsymbol{y}_2=(2,1)^{\mathrm{T}},$$

$$\boldsymbol{d}_2=\boldsymbol{y}_2-\boldsymbol{y}_0=(0,0)^{\mathrm{T}}.$$

因为 $\|\boldsymbol{d}_2\|=0<\varepsilon$,所以迭代终止,得问题(8.1.3)的最优解为 $\boldsymbol{x}_3=\boldsymbol{y}_2=(2,1)^{\mathrm{T}}$. □

顺便指出,Powell 法的进一步改善后来为许多人研究,例如,有人提出把条件(8.4.18)式修改为

$$f(2\boldsymbol{y}_n-\boldsymbol{y}_0)<f(\boldsymbol{y}_0),$$

或

$$[f(\boldsymbol{y}_0)-2f(\boldsymbol{y}_n)+f(2\boldsymbol{y}_n-\boldsymbol{y}_0)]\cdot[f(\boldsymbol{y}_0)-f(\boldsymbol{y}_n)-f(\boldsymbol{y}_m)+f(\boldsymbol{y}_{m+1})]^2$$
$$<\frac{1}{2}[f(\boldsymbol{y}_0)-f(2\boldsymbol{y}_n-\boldsymbol{y}_0)]^2\cdot[f(\boldsymbol{y}_m)-f(\boldsymbol{y}_{m+1})].$$

而且,Powell 法是公认的最有效的算法,至少在目前是这样的. 在一定条件下,Powell 法是收敛的(参见文献[48]).

8.5 单纯形调优法

对于无约束最优化问题(8.1.1),Spendly,Hexit 和 Himsworth(1962)提出一种利用不断调整$\mathbb{R}^n$ 中正则单纯形的顶点来逼近最优解的所谓正则单纯形调优法(regular simplex evolutionary technique). 后来,Nelder 和 Mead(1965)对上述方法做了改进,使之成为一个相当实用的直接法,称之为单纯形调优法(simplex evolutionary technique).

8.5.1 正则单纯形调优法

给定一个初始点 $\boldsymbol{x}_0=(x_1^{(0)},x_2^{(0)},\cdots,x_n^{(0)})^{\mathrm{T}}$ 及数 $t>0$,记

$$a=\frac{t}{n\sqrt{2}}(\sqrt{n+1}+n-1),b=\frac{t}{n\sqrt{2}}(\sqrt{n+1}-1),\qquad(8.5.1)$$

取点 $\boldsymbol{x}_k=(x_1^{(k)},x_2^{(k)},\cdots,x_n^{(k)})^{\mathrm{T}}(k=1,2,\cdots,n)$ 如下：

$$x_j^{(k)}=\begin{cases}x_j^{(0)}+b, j\neq k,\\ x_j^{(0)}+a, j=k.\end{cases} \tag{8.5.2}$$

不难验证，(8.5.1)和(8.5.2)两式确定的多面体 $H(\{\boldsymbol{x}_0,\boldsymbol{x}_1,\cdots,\boldsymbol{x}_n\})$ 是一个具有 $n+1$ 个顶点且棱长为 t 的正则单纯形.

比较上述正则单纯形 $H(\{\boldsymbol{x}_0,\boldsymbol{x}_1,\cdots,\boldsymbol{x}_n\})$ 中各顶点处的目标函数值，通过对顶点重新编号，可使得

$$f(\boldsymbol{x}_0)\leqslant f(\boldsymbol{x}_1)\leqslant\cdots\leqslant f(\boldsymbol{x}_{n-1})\leqslant f(\boldsymbol{x}_n),$$

并称 $\boldsymbol{x}_0$ 为 $H(\{\boldsymbol{x}_0,\boldsymbol{x}_1,\cdots,\boldsymbol{x}_n\})$ 中的最好顶点，$\boldsymbol{x}_n$ 为最坏顶点，$\boldsymbol{x}_{n-1}$ 为次坏顶点.

正则单纯形调优法的基本思想是不断用好的点去替换最坏顶点. 实现这种思想主要是采用下面两种基本技巧.

(1) 反映

首先，求 $H(\{\boldsymbol{x}_0,\boldsymbol{x}_1,\cdots,\boldsymbol{x}_n\})$ 中去掉最坏顶点 $\boldsymbol{x}_n$ 以后得到的 $\mathbb{R}^n$ 中具有 n 个顶点的单纯形 $H(\{\boldsymbol{x}_0,\boldsymbol{x}_1,\cdots,\boldsymbol{x}_{n-1}\})$ 的重心

$$\boldsymbol{x}_{n+1}=\frac{1}{n}\sum_{j=0}^{n-1}\boldsymbol{x}_j,$$

再求最坏顶点 $\boldsymbol{x}_n$ 关于重心 $\boldsymbol{x}_{n+1}$ 的反映点

$$\boldsymbol{x}_{n+2}=\boldsymbol{x}_{n+1}+(\boldsymbol{x}_{n+1}-\boldsymbol{x}_n)=2\boldsymbol{x}_{n+1}-\boldsymbol{x}_n. \tag{8.5.3}$$

图 8.5.1 画出了 $n=2$ 的情形.

如果 $f(\boldsymbol{x}_{n+2})<f(\boldsymbol{x}_{n-1})$，即反映点比次坏顶点好，则用 $\boldsymbol{x}_{n+2}$ 替换 $\boldsymbol{x}_n$，构成新的正则单纯形 $H(\{\boldsymbol{x}_0,\boldsymbol{x}_1,\cdots,\boldsymbol{x}_{n-1},\boldsymbol{x}_{n+2}\})$，称之为反映单纯形. 此时 $\boldsymbol{x}_{n+2}$ 不会成为反映单纯形的最坏顶点. 图 8.5.2，右上方带阴影的三角形即为 $n=2$ 时的反映单纯形.

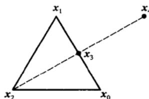


图 8.5.1　反映

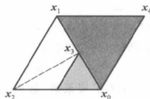


图 8.5.2　紧缩

(2) 紧缩

如果 $f(\boldsymbol{x}_{n+2})\geqslant f(\boldsymbol{x}_{n-1})$，考虑到最好顶点 $\boldsymbol{x}_0$ 可能接近最优解，则将正则单纯形同顶点 $\boldsymbol{x}_0$ 处紧缩，这时新的正则单纯形（称为紧缩单纯形）的顶点取为

$$\boldsymbol{x}_j:=\boldsymbol{x}_0+\frac{1}{2}(\boldsymbol{x}_j-\boldsymbol{x}_0), j=0,1,\cdots,n. \tag{8.5.4}$$

图 8.5.2 左下方阴影的小三角形即为 $n=2$ 时的紧缩单纯形.

由上述的反映和紧缩技巧可以发现,正则单纯形调优法中出现的单纯形都是正则的.

正则单纯形调优法给人的最大启示是:不必局限于用点列去逼近问题的最优解,可以用一系列正则单纯形去接近最优解.由于它的每次迭代都使用了多个点的目标函数值信息,因此有利于单纯形向最优解方向的逼近,迭代效果是稳定的.但是这个方法仅仅运用了较机械的反映和紧缩两种技巧,有时会因紧缩次数过多而影响它向更广泛范围内搜索最优解,以致单纯形过早在非最优解处紧缩为一个点,从而导致搜索失败.

8.5.2 单纯形调优法

为了防止正则单纯形调优法的搜索失败,并加快收敛速度,Nelder 和 Mead 对正则单纯形调优法作了一些改进:使得迭代中出现的单纯形不一定是正则的,并且单纯形在迭代过程中既能变小也能变大.具体说来,单纯形调优法把正则单纯形调优法中的反映和紧缩两种技巧作了推广,且增加了扩展和收缩两种技巧.

(1) 反映

把(8.5.3)式推广为

$$\boldsymbol{x}_{n+2}=\boldsymbol{x}_{n+1}+\alpha(\boldsymbol{x}_{n+1}-\boldsymbol{x}_n),$$

其中 $\alpha\geqslant 1$ 称为反映系数.

(2) 紧缩

把(8.5.4)式推广为

$$\boldsymbol{x}_j:=\boldsymbol{x}_0+\theta(\boldsymbol{x}_j-\boldsymbol{x}_0),j=0,1,\cdots,n,$$

其中 $\theta\in(0,1)$ 称为紧缩系数.

(3) 扩展

当 $f(\boldsymbol{x}_{n+2})<f(\boldsymbol{x}_0)$,即反映点比最好顶点还好时,方向 $\boldsymbol{x}_{n+2}-\boldsymbol{x}_{n+1}$ 是使目标函数值下降的有利方向,故可在 $\boldsymbol{x}_{n+2}$ 和 $\boldsymbol{x}_{n+1}$ 连线的延长线上,求一个扩展点

$$\boldsymbol{x}_{n+3}=\boldsymbol{x}_{n+1}+\gamma(\boldsymbol{x}_{n+2}-\boldsymbol{x}_{n+1}),$$

其中 $\gamma>1$ 称为扩展系数.如果 $f(\boldsymbol{x}_{n+3})<f(\boldsymbol{x}_0)$,即扩展点比最好顶点要好时,就以 $\boldsymbol{x}_{n+3}$ 替换 $\boldsymbol{x}_n$,构成新的单纯形 $H(\{\boldsymbol{x}_0,\boldsymbol{x}_1,\cdots,\boldsymbol{x}_{n-1},\boldsymbol{x}_{n+3}\})$,称之为扩展单纯形.

(4) 收缩

当 $f(\boldsymbol{x}_{n+2})\geqslant f(\boldsymbol{x}_{n-1})$ 时,为方便起见,将 $\boldsymbol{x}_{n+2}$ 和 $\boldsymbol{x}_n$ 之间目标函数值较小的点记为 $\boldsymbol{x}_n$,另一个记为 $\boldsymbol{x}_{n+2}$,并在点 $\boldsymbol{x}_n$ 与 $\boldsymbol{x}_{n+1}$ 的连线上靠近点 $\boldsymbol{x}_n$ 处求一个收缩点

$$\boldsymbol{x}_{n+4}=\boldsymbol{x}_{n+1}+\beta(\boldsymbol{x}_n-\boldsymbol{x}_{n+1}),$$

其中 $\beta\in(0,1)$ 称为收缩系数.如果 $f(\boldsymbol{x}_{n+4})<f(\boldsymbol{x}_n)$,即收缩点比最坏顶点好,则

以 $\boldsymbol{x}_{n+4}$ 替换 $\boldsymbol{x}_n$，构成新的单纯形 $H(\{\boldsymbol{x}_0,\boldsymbol{x}_1,\cdots,\boldsymbol{x}_{n-1},\boldsymbol{x}_{n+4}\})$，称之为收缩单纯形.

算法 8-6(单纯形调优法)

Step 1　选取初始数据. 选取以 $\boldsymbol{x}_0,\boldsymbol{x}_1,\cdots,\boldsymbol{x}_n$ 为顶点的初始单纯形 $H(\{\boldsymbol{x}_0,\boldsymbol{x}_1,\cdots,\boldsymbol{x}_n\})$，给定反映系数 $\alpha\geqslant 1$，紧缩系数 $0<\theta<1$，扩展系数 $\gamma>1$，收缩系数 $0<\beta<1$，允许误差 $\varepsilon>0$.

Step 2　顶点重新编号. 将单纯形的 $n+1$ 个顶点按目标函数值大小重新编号，使其满足

$$f(\boldsymbol{x}_0)\leqslant f(\boldsymbol{x}_1)\leqslant\cdots\leqslant f(\boldsymbol{x}_{n-1})\leqslant f(\boldsymbol{x}_n).$$

Step 3　检查是否满足终止准则. 求单纯形 $H(\{\boldsymbol{x}_0,\boldsymbol{x}_1,\cdots,\boldsymbol{x}_{n-1}\})$ 的重心

$$\boldsymbol{x}_{n+1}=\frac{1}{n}\sum_{j=0}^{n-1}\boldsymbol{x}_j.$$

若 $\left\{\frac{1}{n+1}\sum_{j=0}^{n}[f(\boldsymbol{x}_j)-f(\boldsymbol{x}_{n+1})]^2\right\}^{\frac{1}{2}}\leqslant\varepsilon$，迭代终止，$\boldsymbol{x}_0$ 为问题(8.1.1)的近似最优解；否则，转 Step 4.

Step 4　求反映点. 计算反映点

$$\boldsymbol{x}_{n+2}=\boldsymbol{x}_{n+1}+\alpha(\boldsymbol{x}_{n+1}-\boldsymbol{x}_n).$$

若 $f(\boldsymbol{x}_{n+2})<f(\boldsymbol{x}_0)$，转 Step 5；否则，当 $f(\boldsymbol{x}_{n+2})<f(\boldsymbol{x}_{n-1})$ 时转 Step 6，当 $f(\boldsymbol{x}_{n+2})\geqslant f(\boldsymbol{x}_{n-1})$ 时转 Step 7.

Step 5　扩展. 计算扩展点

$$\boldsymbol{x}_{n+3}=\boldsymbol{x}_{n+1}+\gamma(\boldsymbol{x}_{n+2}-\boldsymbol{x}_{n+1}).$$

若 $f(\boldsymbol{x}_{n+3})<f(\boldsymbol{x}_0)$，令 $\boldsymbol{x}_n:=\boldsymbol{x}_{n+3}$，构造扩展单纯形，转 Step 2；否则，转 Step 6.

Step 6　反映. 令 $\boldsymbol{x}_n:=\boldsymbol{x}_{n+2}$，构造反映单纯形，返回 Step 2.

Step 7　收缩. 把 $\boldsymbol{x}_n$ 和 $\boldsymbol{x}_{n+2}$ 之间对应目标函数值较小的点记作 $\boldsymbol{x}_n$，计算收缩点

$$\boldsymbol{x}_{n+4}=\boldsymbol{x}_{n+1}+\beta(\boldsymbol{x}_n-\boldsymbol{x}_{n+1}).$$

若 $f(\boldsymbol{x}_{n+4})<f(\boldsymbol{x}_n)$，令 $\boldsymbol{x}_n:=\boldsymbol{x}_{n+4}$，构造收缩单纯形，返回 Step 2；否则，转 Step 8.

Step 8　紧缩. 令

$$\boldsymbol{x}_j:=\boldsymbol{x}_0+\theta(\boldsymbol{x}_j-\boldsymbol{x}_0),\ j=0,1,\cdots,n,$$

构造紧缩单纯形，返回 Step 2.

例 8.5.1　用单纯形调优法求解问题(8.1.3)，取初始单纯形

$$H(\{\boldsymbol{x}_0=(0,0)^{\mathrm{T}},\boldsymbol{x}_1=(-2,0)^{\mathrm{T}},\boldsymbol{x}_2=(0,2)^{\mathrm{T}}\})$$

给定 $\alpha=1,\theta=0.5,\gamma=2,\beta=0.5,\varepsilon=0.35$.

解　第一次迭代：

因为

$$f(\boldsymbol{x}_0)=0, f(\boldsymbol{x}_1)=10, f(\boldsymbol{x}_2)=4,$$

把初始单纯形的顶点重新编号,得

$$\boldsymbol{x}_0=(0,0)^{\mathrm{T}}, f(\boldsymbol{x}_0)=0,$$

$$\boldsymbol{x}_1=(0,2)^{\mathrm{T}}, f(\boldsymbol{x}_1)=4,$$

$$\boldsymbol{x}_2=(-2,0)^{\mathrm{T}}, f(\boldsymbol{x}_2)=10.$$

求 $H\{(\boldsymbol{x}_0,\boldsymbol{x}_1)\}$ 的重心

$$\boldsymbol{x}_3=\frac{1}{2}(\boldsymbol{x}_0+\boldsymbol{x}_1)=(0,1)^{\mathrm{T}}.$$

由于 $f(\boldsymbol{x}_3)=1$,

$$\left\{\frac{1}{3}\sum_{j=0}^{2}[f(\boldsymbol{x}_j)-f(\boldsymbol{x}_3)]^2\right\}^{\frac{1}{2}}=5.507\ 571>\varepsilon,$$

因此,再求反映点

$$\boldsymbol{x}_4=\boldsymbol{x}_3+(\boldsymbol{x}_3-\boldsymbol{x}_2)=(2,2)^{\mathrm{T}}.$$

因为 $f(\boldsymbol{x}_4)=-2<f(\boldsymbol{x}_0)$,所以求扩展点

$$\boldsymbol{x}_5=\boldsymbol{x}_3+2(\boldsymbol{x}_4-\boldsymbol{x}_3)=(4,3)^{\mathrm{T}}.$$

又 $f(\boldsymbol{x}_5)=1\geqslant f(\boldsymbol{x}_0)$,故令 $\boldsymbol{x}_2:=\boldsymbol{x}_4$,构造反映单纯形

$$H(\{\boldsymbol{x}_0=(0,0)^{\mathrm{T}},\boldsymbol{x}_1=(0,2)^{\mathrm{T}},\boldsymbol{x}_2=(2,2)^{\mathrm{T}}\}).$$

第二次迭代:

因为

$$f(\boldsymbol{x}_0)=0, f(\boldsymbol{x}_1)=4, f(\boldsymbol{x}_2)=-2,$$

所以把单纯形的顶点重新编号,得

$$\boldsymbol{x}_0=(2,2)^{\mathrm{T}}, f(\boldsymbol{x}_0)=-2,$$

$$\boldsymbol{x}_1=(0,0)^{\mathrm{T}}, f(\boldsymbol{x}_1)=0,$$

$$\boldsymbol{x}_2=(0,2)^{\mathrm{T}}, f(\boldsymbol{x}_2)=4.$$

求 $H(\{\boldsymbol{x}_0,\boldsymbol{x}_1\})$ 的重心

$$\boldsymbol{x}_3=\frac{1}{2}(\boldsymbol{x}_0+\boldsymbol{x}_1)=(1,1)^{\mathrm{T}}.$$

由于 $f(\boldsymbol{x}_3)=-2$,

$$\left\{\frac{1}{3}\sum_{j=0}^{2}[f(\boldsymbol{x}_j)-f(\boldsymbol{x}_3)]^2\right\}^{\frac{1}{2}}=3.651\ 484>\varepsilon,$$

因此再求反映点

$$\boldsymbol{x}_4=\boldsymbol{x}_3+(\boldsymbol{x}_3-\boldsymbol{x}_2)=(2,0)^{\mathrm{T}}.$$

因为 $f(\boldsymbol{x}_4)=-2\geqslant f(\boldsymbol{x}_0)$, $f(\boldsymbol{x}_4)<f(\boldsymbol{x}_1)$,所以令 $\boldsymbol{x}_2:=\boldsymbol{x}_4$,构造反映单纯形

$$H(\{\boldsymbol{x}_0=(2,2)^{\mathrm{T}},\boldsymbol{x}_1=(0,0)^{\mathrm{T}},\boldsymbol{x}_2=(2,0)^{\mathrm{T}}\}).$$

第三次迭代：

因为

$$f(\boldsymbol{x}_0) = -2, f(\boldsymbol{x}_1) = 0, f(\boldsymbol{x}_2) = -2,$$

所以把单纯形的顶点重新编号为

$$\boldsymbol{x}_0 = (2,2)^{\mathrm{T}}, f(\boldsymbol{x}_0) = -2,$$

$$\boldsymbol{x}_1 = (2,0)^{\mathrm{T}}, f(\boldsymbol{x}_1) = -2,$$

$$\boldsymbol{x}_2 = (0,0)^{\mathrm{T}}, f(\boldsymbol{x}_2) = 0.$$

求 $H(\{\boldsymbol{x}_0, \boldsymbol{x}_1\})$ 的重心

$$\boldsymbol{x}_3 = \frac{1}{2}(\boldsymbol{x}_0 + \boldsymbol{x}_1) = (2,1)^{\mathrm{T}}.$$

由于 $f(\boldsymbol{x}_3) = -3$，

$$\left\{\frac{1}{3}\sum_{j=0}^{2}[f(\boldsymbol{x}_j) - f(\boldsymbol{x}_3)]^2\right\}^{\frac{1}{2}} = 1.914\,854 > \varepsilon,$$

因此再求反映点

$$\boldsymbol{x}_4 = \boldsymbol{x}_3 + (\boldsymbol{x}_3 - \boldsymbol{x}_2) = (4,2)^{\mathrm{T}}.$$

因为 $f(\boldsymbol{x}_4) = 0 \geqslant f(\boldsymbol{x}_0), f(\boldsymbol{x}_4) \geqslant f(\boldsymbol{x}_1)$，所以把 $\boldsymbol{x}_4$ 和 $\boldsymbol{x}_2$ 中对应目标函数值较小的点记作 $\boldsymbol{x}_2$，求收缩点

$$\boldsymbol{x}_6 = \boldsymbol{x}_3 + 0.5(\boldsymbol{x}_2 - \boldsymbol{x}_3) = (1, 0.5)^{\mathrm{T}}.$$

因 $f(\boldsymbol{x}_6) = -2.25 < f(\boldsymbol{x}_2)$，故令 $\boldsymbol{x}_2 := \boldsymbol{x}_6$，构造收缩单纯形

$$H(\{\boldsymbol{x}_0 = (2,2)^{\mathrm{T}}, \boldsymbol{x}_1 = (2,0)^{\mathrm{T}}, \boldsymbol{x}_2 = (1,0.5)^{\mathrm{T}}\})$$

第四次迭代：

因为

$$f(\boldsymbol{x}_0) = -2, f(\boldsymbol{x}_1) = -2, f(\boldsymbol{x}_2) = -2.25,$$

所以把单纯形的顶点重新编号为

$$\boldsymbol{x}_0 = (1,0.5)^{\mathrm{T}}, f(\boldsymbol{x}_0) = -2.25,$$

$$\boldsymbol{x}_1 = (2,2)^{\mathrm{T}}, f(\boldsymbol{x}_1) = -2,$$

$$\boldsymbol{x}_2 = (2,0)^{\mathrm{T}}, f(\boldsymbol{x}_2) = -2.$$

求 $H(\{\boldsymbol{x}_0, \boldsymbol{x}_1\})$ 的重心

$$\boldsymbol{x}_3 = \frac{1}{2}(\boldsymbol{x}_0 + \boldsymbol{x}_1) = (1.5, 1.25)^{\mathrm{T}}.$$

由于 $f(\boldsymbol{x}_3) = -2.562\,5$，

$$\left\{\frac{1}{3}\sum_{j=0}^{2}[f(\boldsymbol{x}_j) - f(\boldsymbol{x}_3)]^2\right\}^{\frac{1}{2}} = 0.493\,447 > \varepsilon,$$

因此，再求反映点

$$x_4 = x_3 + (x_3 - x_2) = (1,2.5)^T.$$

因为$f(x_4)=1.75 \geqslant f(x_0)$，$f(x_4) \geqslant f(x_1)$，所以把 x_4 和 x_2 中对应目标函数值较小的点记作 x_2，求收缩点

$$x_6 = x_3 + 0.5(x_2 - x_3) = (1.25,1.875)^T.$$

因$f(x_6) = -1.05625 \geqslant f(x_2)$，故令

$$x_0 := x_0 + 0.5(x_0 - x_0) = (1,0.5)^T,$$

$$x_1 := x_0 + 0.5(x_1 - x_0) = (1.5,1.25)^T,$$

$$x_2 := x_0 + 0.5(x_2 - x_0) = (1.5,0.25)^T.$$

构造紧缩单纯形

$$H(\{x_0 = (1,0.5)^T, x_1 = (1.5,1.25)^T, x_2 = (1.5,0.25)^T\}).$$

第五次迭代：

因为

$$f(x_0) = -2.25, f(x_1) = -2.5625, f(x_2) = -2.5625,$$

所以，单纯形的顶点重新编号为

$$x_0 = (1.5,1.25)^T, f(x_0) = -2.5625,$$

$$x_1 = (1.5,0.25)^T, f(x_1) = -2.5625,$$

$$x_2 = (1,0.5)^T, f(x_2) = -2.25.$$

求 $H(\{x_0,x_1\})$的重心

$$x_3 = \frac{1}{2}(x_0 + x_1) = (1.5,0.75)^T.$$

由于$f(x_3) = -2.8125$，

$$\left\{\frac{1}{3}\sum_{j=0}^{2}[f(x_j) - f(x_3)]^2\right\}^{\frac{1}{2}} = 0.383582 > \varepsilon,$$

因此，求反映点

$$x_4 = x_3 + (x_3 - x_2) = (2,1)^T.$$

因为$f(x_4) = -3 < f(x_0)$，所以求扩展点

$$x_5 = x_3 + 2(x_4 - x_3) = (2.5,1.25)^T.$$

因$f(x_5) = -2.825 < f(x_0)$，故令 $x_2 := x_5$，构造扩展单纯形

$$H(\{x_0 = (1.5,1.25)^T, x_1 = (1.5,0.25)^T, x_2 = (2.5,1.25)^T\}).$$

第六次迭代：

因为

$$f(x_0) = -2.5625, f(x_1) = -2.5625, f(x_2) = -2.8125,$$

所以单纯形的顶点重新编号为

$$\boldsymbol{x}_0=(2.5,1.25)^{\mathrm{T}}, f(\boldsymbol{x}_0)=-2.8125,$$
$$\boldsymbol{x}_1=(1.5,1.25)^{\mathrm{T}}, f(\boldsymbol{x}_1)=-2.5625,$$
$$\boldsymbol{x}_2=(1.5,0.25)^{\mathrm{T}}, f(\boldsymbol{x}_2)=-2.5625.$$

求 $H(\{\boldsymbol{x}_0,\boldsymbol{x}_1\})$ 的重心

$$\boldsymbol{x}_3=\frac{1}{2}(\boldsymbol{x}_0+\boldsymbol{x}_1)=(2,1.25)^{\mathrm{T}}.$$

由于 $f(\boldsymbol{x}_3)=-2.9375$,

$$\left\{\frac{1}{3}\sum_{j=0}^{2}[f(\boldsymbol{x}_j)-f(\boldsymbol{x}_3)]^2\right\}^{\frac{1}{2}}=0.314576<\varepsilon,$$

因此迭代终止,得问题(8.1.3)的近似最优解 $\boldsymbol{x}_0=(2.5,1.25)^{\mathrm{T}}$. □

在适当条件下,正则单纯形调优法是收敛的(参见文献[16,49]). 单纯形调优法虽然求解速度比正则单纯形调优法要快而且有效,但至今仍处于经验方法阶段. 单纯形调优法对于变量不多且精度要求不高的问题的求解是很有用的,但是当变量个数较多时,这个方法的效果一般是不好的.

习 题 八

1. 用坐标轮换法求解问题:

(1) $\min 3x_1^2+5x_2^2$,

取初始点 $\boldsymbol{x}_0=(2,1)^{\mathrm{T}}$,允许误差 $\varepsilon=0.1$;

(2) $\min x_1^2+x_2^2-4x_1-2x_1x_2$,

取初始点 $\boldsymbol{x}_0=(1,1)^{\mathrm{T}}$,允许误差 $\varepsilon=0.1$.

2. 用模式搜索法求解第1题,其中均取初始步长 $\delta=1$,收缩因子 $\alpha=0.5$.

3. 用旋转方向法求解下列问题,均取初始搜索方向组为两个坐标轴方向 $\boldsymbol{e}_1$ 和 $\boldsymbol{e}_2$,初始步长 $\boldsymbol{\delta}_0=(1,1)^{\mathrm{T}}$,收缩因子 $\alpha=0.5$,放大因子 $\beta=3$,要求迭代两次.

(1) $\min(x_1-3)^2+2(x_2+2)^2$,

取初始点 $\boldsymbol{x}_0=(0,0)^{\mathrm{T}}$;

(2) $\min 3x_1^2+x_2^2-2x_1x_2+4x_1+3x_2$,

取初始点 $\boldsymbol{x}_0=(4,5)^{\mathrm{T}}$.

4. 设对一切 $j=1,2,\cdots,n$,有

$$\boldsymbol{p}_j=\sum_{i=j}^{n}\lambda_i\boldsymbol{d}_i,\lambda_i\neq 0,\boldsymbol{d}_i\in\mathbb{R}^n.$$

若 $\boldsymbol{d}_1,\boldsymbol{d}_2,\cdots,\boldsymbol{d}_n$ 线性无关,试证 $\boldsymbol{p}_1,\boldsymbol{p}_2,\cdots,\boldsymbol{p}_n$ 线性无关.

5. 用原始 Powell 法求解问题

$$\min \frac{3}{2}x_1^2+\frac{1}{2}x_2^2-x_1x_2-2x_1,$$

取初始点 $\boldsymbol{x}_0=(-2,4)^{\mathrm{T}}$,初始搜索方向组为两个坐标轴方向 $\boldsymbol{e}_1$ 和 $\boldsymbol{e}_2$.

6. 用 Powell 法求解问题

$$\min\ 2x_1^2 + x_2^2 - x_1 x_3 - 4x_2 ,$$

取初始点 $\boldsymbol{x}_0 = (1,1)^{\mathrm{T}}$,初始搜索方向组为两个坐标轴方向 $\boldsymbol{e}_1$ 和 $\boldsymbol{e}_2$,允许误差 $\varepsilon = 10^{-6}$.

7. 设 $\boldsymbol{A} = (a_{ij}) \in \mathbb{R}^{n\times n}$为半正定矩阵,证明 Hadamard 不等式

$$\det \boldsymbol{A} \leqslant a_{11} a_{22} \cdots a_{nn} ,$$

并且上式中等号成立的充要条件是 $\boldsymbol{A}$ 为对角矩阵.

8. 证明:由(8.5.1)和(8.5.2)两式确定的多面体 $H(\{\boldsymbol{x}_0, \boldsymbol{x}_1, \cdots, \boldsymbol{x}_n\})$是一个具有 $n+1$ 个顶点且棱长为 t 的正则单纯形.

9. 用单纯形调优法求解问题

$$\min\ 4(x_1 - 5)^2 + (x_2 - 6)^2 ,$$

取初始单纯形为

$$H(\{\boldsymbol{x}_0 = (8,9)^{\mathrm{T}}, \boldsymbol{x}_1 = (10,11)^{\mathrm{T}}, \boldsymbol{x}_2 = (8,11)^{\mathrm{T}}\}).$$

给定反映系数 $\alpha = 1$,紧缩系数 $\theta = 0.5$,扩展系数 $\gamma = 2$,收缩系数 $\beta = 0.5$,允许误差 $\varepsilon = 6$.

第九章　可行方向法

从本章开始，我们着手讨论求解约束非线性规划问题的方法，即约束最优化方法. 可行方向法(feasible directions methods)是其中的一类算法. 此类方法可看做无约束下降算法的自然推广，其基本思想是从可行点出发，沿可行下降方向进行搜索，求出使目标函数值下降的新的可行点. 算法包括选择搜索方向和确定搜索步长两个主要方面. 搜索方向的选择方式不同就形成不同的可行方向法.

9.1　Zoutendijk 可行方向法

Zoutendijk 可行方向法包括利用起作用约束构造可行方向和 ε 起作用约束构造可行方向.

9.1.1　线性约束情形

考虑线性约束问题

$$\begin{cases} \min & f(\boldsymbol{x}); \\ \text{s.t.} & \boldsymbol{A}\boldsymbol{x} \geqslant \boldsymbol{b}, \\ & \boldsymbol{E}\boldsymbol{x} = \boldsymbol{e}, \end{cases} \tag{9.1.1}$$

其中 $f(\boldsymbol{x})$ 为可微函数，$\boldsymbol{A}$ 为 $m \times n$ 矩阵，$\boldsymbol{E}$ 为 $l \times n$ 矩阵，$\boldsymbol{x} \in \mathbb{R}^n$，$\boldsymbol{b} \in \mathbb{R}^m$，$\boldsymbol{e} \in \mathbb{R}^l$.

(1) 利用起作用约束构造可行下降方向

定理 9.1.1　设 $\boldsymbol{x}$ 是问题(9.1.1)的可行解，在点 $\boldsymbol{x}$ 处有 $\boldsymbol{A}_1\boldsymbol{x} = \boldsymbol{b}_1$，$\boldsymbol{A}_2\boldsymbol{x} > \boldsymbol{b}_2$，其中

$$\boldsymbol{A} = \begin{bmatrix} \boldsymbol{A}_1 \\ \boldsymbol{A}_2 \end{bmatrix}, \boldsymbol{b} = \begin{bmatrix} \boldsymbol{b}_1 \\ \boldsymbol{b}_2 \end{bmatrix},$$

则非零向量 $\boldsymbol{d}$ 为 $\boldsymbol{x}$ 处的可行方向的充要条件是

$$\boldsymbol{A}_1\boldsymbol{d} \geqslant \boldsymbol{0}, \boldsymbol{E}\boldsymbol{d} = \boldsymbol{0}. \tag{9.1.2}$$

证明　先证必要性. 设非零向量 $\boldsymbol{d}$ 为 $\boldsymbol{x}$ 处的可行方向，则存在数 $\delta > 0$，使得对每个 $\lambda \in (0, \delta)$，有 $\boldsymbol{x} + \lambda\boldsymbol{d}$ 为可行点，即

$$\begin{aligned} \boldsymbol{A}(\boldsymbol{x} + \lambda\boldsymbol{d}) &\geqslant \boldsymbol{b}, \\ \boldsymbol{E}(\boldsymbol{x} + \lambda\boldsymbol{d}) &= \boldsymbol{e}. \end{aligned} \tag{9.1.3}$$

因为

$$\boldsymbol{A}(\boldsymbol{x}+\lambda\boldsymbol{d})=\begin{bmatrix}\boldsymbol{A}_1\boldsymbol{x}+\lambda\boldsymbol{A}_1\boldsymbol{d}\\\boldsymbol{A}_2\boldsymbol{x}+\lambda\boldsymbol{A}_2\boldsymbol{d}\end{bmatrix}=\begin{bmatrix}\boldsymbol{b}_1+\lambda\boldsymbol{A}_1\boldsymbol{d}\\\boldsymbol{A}_2\boldsymbol{x}+\lambda\boldsymbol{A}_2\boldsymbol{d}\end{bmatrix}\geqslant\begin{bmatrix}\boldsymbol{b}_1\\\boldsymbol{b}_2\end{bmatrix},$$

所以 $\boldsymbol{A}_1\boldsymbol{d}\geqslant\boldsymbol{0}$. 并由(9.1.3)式得到 $\boldsymbol{Ed}=\boldsymbol{0}$. 必要性得证.

再证充分性. 设有非零向量 $\boldsymbol{d}$ 满足(9.1.2)式,由 $\boldsymbol{A}_2\boldsymbol{x}>\boldsymbol{b}_2$ 可知,存在充分小的 $\delta>0$,使得对于 $\lambda\in(0,\delta)$,有

$$\boldsymbol{A}_2(\boldsymbol{x}+\lambda\boldsymbol{d})\geqslant\boldsymbol{b}_2. \tag{9.1.4}$$

根据(9.1.2)中第1式及 $\boldsymbol{A}_1\boldsymbol{x}=\boldsymbol{b}_1$ 得知

$$\boldsymbol{A}_1(\boldsymbol{x}+\lambda\boldsymbol{d})\geqslant\boldsymbol{b}_1. \tag{9.1.5}$$

合并(9.1.4)和(9.1.5)两式得到

$$\boldsymbol{A}(\boldsymbol{x}+\lambda\boldsymbol{d})\geqslant\boldsymbol{b}. \tag{9.1.6}$$

再由(9.1.2)中第2式及 $\boldsymbol{Ex}=\boldsymbol{e}$ 可知

$$\boldsymbol{E}(\boldsymbol{x}+\lambda\boldsymbol{d})=\boldsymbol{e}. \tag{9.1.7}$$

这样,(9.1.6)和(9.1.7)两式表明 $\boldsymbol{x}+\lambda\boldsymbol{d}$ 是可行点,因此 $\boldsymbol{d}$ 是 $\boldsymbol{x}$ 处的可行方向.

□

根据定理1.2.3和定理9.1.1,如果非零向量 $\boldsymbol{d}$ 同时满足

$$\nabla f(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{d}<0,\boldsymbol{A}_1\boldsymbol{d}\geqslant\boldsymbol{0},\boldsymbol{Ed}=\boldsymbol{0},$$

则 $\boldsymbol{d}$ 是 $\boldsymbol{x}$ 处的可行下降方向.

因此,确定搜索方向可归结为求解线性规划问题

$$\begin{cases}\min & \nabla f(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{d};\\ \text{s.t.} & \boldsymbol{A}_1\boldsymbol{d}\geqslant\boldsymbol{0},\\ & \boldsymbol{Ed}=\boldsymbol{0},\\ & -1\leqslant d_j\leqslant1,j=1,2,\cdots,n,\end{cases} \tag{9.1.8}$$

其中为了能获得有限最优解,限制了方向 $\boldsymbol{d}$ 的长度,即增加规范约束条件 $-1\leqslant d_j\leqslant1,j=1,2,\cdots,n$.

因 $\boldsymbol{d}=\boldsymbol{0}$ 是问题(9.1.8)的可行解,故目标函数的最优值必定小于或等于0.

若 $\nabla f(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{d}<0$,则 $\boldsymbol{d}$ 为可行下降方向;

若 $\nabla f(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{d}=0$,则下面的定理将证明 $\boldsymbol{x}$ 是K-T点.

定理9.1.2 设 $\boldsymbol{x}$ 是问题(9.1.1)的可行解,在点 $\boldsymbol{x}$ 处有 $\boldsymbol{A}_1\boldsymbol{x}=\boldsymbol{b}_1,\boldsymbol{A}_2\boldsymbol{x}>\boldsymbol{b}_2$,其中

$$\boldsymbol{A}=\begin{bmatrix}\boldsymbol{A}_1\\\boldsymbol{A}_2\end{bmatrix},\boldsymbol{b}=\begin{bmatrix}\boldsymbol{b}_1\\\boldsymbol{b}_2\end{bmatrix},$$

则 $\boldsymbol{x}$ 为K-T点的充要条件是问题(9.1.8)的目标函数最优值为0.

证明 按K-T点的含义,我们知道,$\boldsymbol{x}$ 是问题(9.1.1)的K-T点,即存在

向量 $\boldsymbol{u}$ 和 $\boldsymbol{v}$,$\boldsymbol{u}\geqslant\boldsymbol{0}$,使得

$$\nabla f(\boldsymbol{x})-\boldsymbol{A}_1^{\mathrm{T}}\boldsymbol{u}-\boldsymbol{E}^{\mathrm{T}}\boldsymbol{v}=\boldsymbol{0}. \tag{9.1.9}$$

令 $\boldsymbol{v}=\boldsymbol{p}-\boldsymbol{q}$,$\boldsymbol{p},\boldsymbol{q}\geqslant\boldsymbol{0}$,把(9.1.9)式写成

$$(-\boldsymbol{A}_1^{\mathrm{T}},-\boldsymbol{E}^{\mathrm{T}},\boldsymbol{E}^{\mathrm{T}})\begin{bmatrix}\boldsymbol{u}\\\boldsymbol{p}\\\boldsymbol{q}\end{bmatrix}=-\nabla f(\boldsymbol{x}),\begin{bmatrix}\boldsymbol{u}\\\boldsymbol{p}\\\boldsymbol{q}\end{bmatrix}\geqslant\boldsymbol{0}. \tag{9.1.10}$$

由 Farkas 引理,(9.1.10)式有解的充要条件是

$$\begin{bmatrix}-\boldsymbol{A}_1\\-\boldsymbol{E}\\\boldsymbol{E}\end{bmatrix}\boldsymbol{d}\leqslant\boldsymbol{0},\ -\nabla f(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{d}>0$$

无解,即

$$\nabla f(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{d}<0,\boldsymbol{A}_1\boldsymbol{d}\geqslant\boldsymbol{0},\boldsymbol{E}\boldsymbol{d}=\boldsymbol{0}$$

无解,这等价于问题(9.1.8)的目标函数最优值为 0. □

(2) 确定一维搜索步长

设从可行点 $\boldsymbol{x}_k$ 出发,沿可行下降方向 $\boldsymbol{d}_k$ 作一维搜索,得后继迭代点

$$\boldsymbol{x}_{k+1}=\boldsymbol{x}_k+\lambda_k\boldsymbol{d}_k.$$

为保证后继迭代点的可行性,可考虑求解如下带约束的一维最优化问题来确定步长:

$$\begin{cases}\min & f(\boldsymbol{x}_k+\lambda\boldsymbol{d}_k);\\ \text{s.t.} & \boldsymbol{A}(\boldsymbol{x}_k+\lambda\boldsymbol{d}_k)\geqslant\boldsymbol{b},\\ & \boldsymbol{E}(\boldsymbol{x}_k+\lambda\boldsymbol{d}_k)=\boldsymbol{e},\\ & \lambda\geqslant 0.\end{cases} \tag{9.1.11}$$

由于 $\boldsymbol{d}_k$ 为可行点 $\boldsymbol{x}_k$ 处的可行方向,则

$$\boldsymbol{E}\boldsymbol{x}_k=\boldsymbol{e},\boldsymbol{E}\boldsymbol{d}_k=\boldsymbol{0},$$

因此,问题(9.1.11)中等式约束总是满足的.

将问题(9.1.11)中的第 1 个约束分解出起作用的约束和不起作用的约束.它们对应的系数矩阵分别记作 $\boldsymbol{A}_1$ 和 $\boldsymbol{A}_2$,即

$$\boldsymbol{A}_1\boldsymbol{x}_k=\boldsymbol{b}_1,\boldsymbol{A}_2\boldsymbol{x}_k>\boldsymbol{b}_2 \tag{9.1.12}$$

不妨假设

$$\boldsymbol{A}=\begin{bmatrix}\boldsymbol{A}_1\\\boldsymbol{A}_2\end{bmatrix},\boldsymbol{b}=\begin{bmatrix}\boldsymbol{b}_1\\\boldsymbol{b}_2\end{bmatrix},$$

这样,问题(9.1.11)中的第 1 个约束可以写成

$$\begin{bmatrix}\boldsymbol{A}_1\boldsymbol{x}_k+\lambda\boldsymbol{A}_1\boldsymbol{d}_k\\\boldsymbol{A}_2\boldsymbol{x}_k+\lambda\boldsymbol{A}_2\boldsymbol{d}_k\end{bmatrix}\geqslant\begin{bmatrix}\boldsymbol{b}_1\\\boldsymbol{b}_2\end{bmatrix}. \tag{9.1.13}$$

由于 $\boldsymbol{d}_k$ 为可行方向，则 $\boldsymbol{A}_1\boldsymbol{d}_k \geqslant \boldsymbol{0}$，又 $\lambda \geqslant 0$，$\boldsymbol{A}_1\boldsymbol{x}_k = \boldsymbol{b}_1$，因此在(9.1.13)式中

$$\boldsymbol{A}_1\boldsymbol{x}_k + \lambda\boldsymbol{A}_1\boldsymbol{d}_k \geqslant \boldsymbol{b}_1$$

自然成立. 约束条件(9.1.13)式简化为

$$\boldsymbol{A}_2\boldsymbol{x}_k + \lambda\boldsymbol{A}_2\boldsymbol{d}_k \geqslant \boldsymbol{b}_2,$$

进而问题(9.1.11)简化为

$$\begin{cases} \min & f(\boldsymbol{x}_k + \lambda\boldsymbol{d}_k); \\ \text{s.t.} & \boldsymbol{A}_2\boldsymbol{x}_k + \lambda\boldsymbol{A}_2\boldsymbol{d}_k \geqslant \boldsymbol{b}_2, \\ & \lambda \geqslant 0. \end{cases} \tag{9.1.14}$$

下面对这一问题作进一步简化. 令

$$\hat{\boldsymbol{b}} = \boldsymbol{b}_2 - \boldsymbol{A}_2\boldsymbol{x}_k, \hat{\boldsymbol{d}} = \boldsymbol{A}_2\boldsymbol{d}_k,$$

则由(9.1.12)式中第 2 式知，$\hat{\boldsymbol{b}} < \boldsymbol{0}$. 问题(9.1.14)中的第 1 个约束条件可以写成

$$\lambda\hat{\boldsymbol{d}} \geqslant \hat{\boldsymbol{b}}. \tag{9.1.15}$$

以下分两种情况讨论：

(1) $\hat{\boldsymbol{d}} \geqslant \boldsymbol{0}$，这时对于 $\lambda \geqslant 0$，(9.1.15)式总成立；

(2) $\hat{\boldsymbol{d}} \ngeq \boldsymbol{0}$，这时$\hat{\boldsymbol{d}}$至少有一个分量$\hat{d}_i < 0$. 对于 $\lambda \geqslant 0$，使(9.1.15)式成立的 λ 的上限为 $\min\left\{\frac{\hat{b}_i}{\hat{d}_i} \middle| \hat{d}_i < 0\right\}$. 记

$$\lambda_{\max} = \begin{cases} \min\left\{\frac{\hat{b}_i}{\hat{d}_i} \middle| \hat{d}_i < 0\right\}, & \text{当}\hat{\boldsymbol{d}} \ngeq \boldsymbol{0}\text{ 时}, \\ \infty, & \text{当}\hat{\boldsymbol{d}} \geqslant \boldsymbol{0}\text{ 时}, \end{cases} \tag{9.1.16}$$

则(9.1.14)式的约束条件最终简化为

$$0 \leqslant \lambda \leqslant \lambda_{\max}.$$

于是，问题(9.1.11)可归结为一维搜索

$$\min_{0 \leqslant \lambda \leqslant \lambda_{\max}} f(\boldsymbol{x}_k + \lambda\boldsymbol{d}_k). \tag{9.1.17}$$

综上所述，可以得到求解线性约束问题(9.1.1)的可行方向法的基本思想：先分解出迭代点的起作用约束，通过求解线性规划问题(9.1.8)得到可行下降方向，再由(9.1.17)式进行一维搜索来求得后继迭代点. 这种方法是 Zoutendijk (1960)给出的，叫做 Zoutendijk 法. 下面给出该算法的具体步骤.

算法 9-1(线性约束 Zoutendijk 法)

Step 1 选取初始数据. 选取初始可行点 $\boldsymbol{x}_0$，允许误差 $\varepsilon > 0$，令 $k = 0$.

Step 2 分解出起作用约束. 在 $\boldsymbol{x}_k$ 处分解

$$A=\begin{bmatrix}A_1\\A_2\end{bmatrix},b=\begin{bmatrix}b_1\\b_2\end{bmatrix},$$

使得 $A_1x_k=b_1, A_2x_k>b_2$，计算 $\nabla f(x_k)$.

Step 3　求解辅助线性规划问题. 求解线性规划问题

$$\begin{cases}\min & \nabla f(x_k)^{\mathrm{T}}d;\\ \text{s.t.} & A_1d\geqslant 0,\\ & Ed=0,\\ & -1\leqslant d_j\leqslant 1, j=1,2,\cdots,n,\end{cases}$$

得最优解 d_k.

Step 4　构造可行下降方向. 若 $|\nabla f(x_k)^{\mathrm{T}}d_k|\leqslant\varepsilon$，停止迭代，输出 x_k；否则，得到可行下降方向 d_k，转 Step 5.

Step 5　进行一维搜索. 计算 $\hat{b}=b_2-A_2x_k, \hat{d}=A_2d_k$，按(9.1.16)式求出 $\lambda_{\max}$，然后求解

$$\min_{0\leqslant\lambda\leqslant\lambda_{\max}} f(x_k+\lambda d_k).$$

得最优解 λ_k. 令 $x_{k+1}=x_k+\lambda_k d_k, k:=k+1$，返回 Step 2.

现在，我们来看如何求得问题(9.1.1)的一个初始可行点. 引入人工变量(向量) ξ,η，解辅助线性规划问题

$$\begin{cases}\min & \sum_{i=1}^{m}\xi_i+\sum_{j=1}^{l}\eta_j;\\ \text{s.t.} & Ax+\xi\geqslant b,\\ & Ex+\eta=e,\\ & \xi\geqslant 0,\eta\geqslant 0.\end{cases}\tag{9.1.18}$$

如果问题(9.1.18)的最优解为

$$(\bar{x},\bar{\xi},\bar{\eta})=(\bar{x},0,0),$$

那么 $\bar{x}$ 就是问题(9.1.1)的一个可行解. 当然，如果方便的话，也可通过观察和试算求得初始可行解.

例 9.1.1　用 Zoutendijk 法解下列问题

$$\begin{cases}\min & 2x_1^2+2x_2^2-2x_1x_2-4x_1-6x_2;\\ \text{s.t.} & -x_1-x_2\geqslant -2,\\ & -x_1-5x_2\geqslant -5,\\ & x_1\geqslant 0,\\ & x_2\geqslant 0,\end{cases}\tag{9.1.19}$$

要求选取初始可行点 $x_0=(0,0)^{\mathrm{T}}$.

解 $\nabla f(\boldsymbol{x})=\begin{bmatrix}4x_1-2x_2-4\\4x_2-2x_1-6\end{bmatrix}$.

第一次迭代:

在 $\boldsymbol{x}_0$ 处,起作用约束指标集 $I(\boldsymbol{x}_0)=\{3,4\}$,对应有

$$\boldsymbol{A}_1=\begin{bmatrix}1&0\\0&1\end{bmatrix},\boldsymbol{b}_1=\begin{bmatrix}0\\0\end{bmatrix},\boldsymbol{A}_2=\begin{bmatrix}-1&-1\\-1&-5\end{bmatrix},\boldsymbol{b}_2=\begin{bmatrix}-2\\-5\end{bmatrix},$$

$$\nabla f(\boldsymbol{x}_0)=(-4,-6)^{\mathrm{T}}.$$

求解辅助线性规划问题

$$\begin{cases}\min & -4d_1-6d_2;\\ \text{s.t.} & d_1\geqslant 0,\\ & -1\leqslant d_1\leqslant 1,\\ & -1\leqslant d_2\leqslant 1,\end{cases}$$

得最优解 $\boldsymbol{d}_0=(1,1)^{\mathrm{T}}$.

因 $|\nabla f(\boldsymbol{x}_0)^{\mathrm{T}}\boldsymbol{d}_0|=10\neq 0$,故 $\boldsymbol{d}_0$ 为可行下降方向,沿 $\boldsymbol{d}_0$ 作一维搜索,计算

$$\hat{\boldsymbol{b}}=\boldsymbol{b}_2-\boldsymbol{A}_2\boldsymbol{x}_0=\begin{bmatrix}-2\\-5\end{bmatrix}-\begin{bmatrix}-1&-1\\-1&-5\end{bmatrix}\begin{bmatrix}0\\0\end{bmatrix}=\begin{bmatrix}-2\\-5\end{bmatrix},$$

$$\hat{\boldsymbol{d}}=\boldsymbol{A}_2\boldsymbol{d}_0=\begin{bmatrix}-1&-1\\-1&-5\end{bmatrix}\begin{bmatrix}1\\1\end{bmatrix}=\begin{bmatrix}-2\\-6\end{bmatrix},$$

$$\lambda_{\max}=\min\left\{\frac{-2}{-2},\frac{-5}{-6}\right\}=\frac{5}{6}.$$

求解

$$\min_{0\leqslant\lambda\leqslant\frac{5}{6}} f(\boldsymbol{x}_0+\lambda\boldsymbol{d}_0)=2\lambda^2-10\lambda,$$

得到 $\lambda_0=\dfrac{5}{6}$,所以

$$\boldsymbol{x}_1=\boldsymbol{x}_0+\lambda_0\boldsymbol{d}_0=\begin{bmatrix}0\\0\end{bmatrix}+\frac{5}{6}\begin{bmatrix}1\\1\end{bmatrix}=\begin{bmatrix}\frac{5}{6}\\\frac{5}{6}\end{bmatrix}.$$

第二次迭代:

$$I(\boldsymbol{x}_1)=\{2\},$$

$$\boldsymbol{A}_1=(-1,-5),\boldsymbol{b}_1=(-5),\boldsymbol{A}_2=\begin{bmatrix}-1&-1\\1&0\\0&1\end{bmatrix},\boldsymbol{b}_2=\begin{bmatrix}-2\\0\\0\end{bmatrix},$$

$$\nabla f(\boldsymbol{x}_1)=\left(-\frac{7}{3},-\frac{13}{3}\right)^{\mathrm{T}}.$$

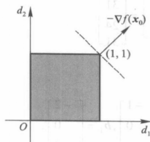


图 9.1.1 第一次迭代的图解法

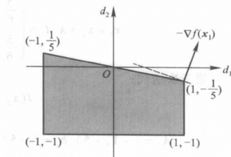


图 9.1.2 第二次迭代的图解法

求解辅助线性规划问题

$$\begin{cases} \min & -\dfrac{7}{3}d_1-\dfrac{13}{3}d_2; \\ \text{s.t.} & -d_1-5d_2\geqslant 0, \\ & -1\leqslant d_1\leqslant 1, \\ & -1\leqslant d_2\leqslant 1, \end{cases}$$

得最优解 $d_1=\left(1,-\dfrac{1}{5}\right)^{\mathrm{T}}$.

因 $|\nabla f(\boldsymbol{x}_0)^{\mathrm{T}}\boldsymbol{d}_1|=\dfrac{22}{15}\neq 0$,故 $\boldsymbol{d}_1$ 为可行下降方向,沿 $\boldsymbol{d}_1$ 作一维搜索,计算

$$\hat{\boldsymbol{b}}=\boldsymbol{b}_2-\boldsymbol{A}_2\boldsymbol{x}_1=\begin{bmatrix}-2\\0\\0\end{bmatrix}-\begin{bmatrix}-1&-1\\1&0\\0&1\end{bmatrix}\begin{bmatrix}5/6\\5/6\end{bmatrix}=\begin{bmatrix}-1/3\\-5/6\\-5/6\end{bmatrix},$$

$$\hat{\boldsymbol{d}}=\boldsymbol{A}_2\boldsymbol{d}_1=\begin{bmatrix}-1&-1\\1&0\\0&1\end{bmatrix}\begin{bmatrix}1\\-\dfrac{1}{5}\end{bmatrix}=\left(-\frac{4}{5},1,-\frac{1}{5}\right)^{\mathrm{T}},$$

$$\lambda_{\max}=\min\left\{\frac{-1/3}{-4/5},\frac{-5/6}{-1/5}\right\}=\frac{5}{12}.$$

求解

$$\min_{0\leqslant\lambda\leqslant\frac{5}{12}}f(\boldsymbol{x}_1+\lambda\boldsymbol{d}_1)=\frac{62}{25}\lambda^2-\frac{22}{15}\lambda,$$

得 $\lambda_1=\dfrac{55}{186}$,所以

$$x_2 = x_1 + \lambda_1 d_1 = \begin{bmatrix} \frac{5}{6} \\ \frac{5}{6} \end{bmatrix} + \frac{55}{186}\begin{bmatrix} 1 \\ -\frac{1}{5} \end{bmatrix} = \begin{bmatrix} \frac{35}{31} \\ \frac{24}{31} \end{bmatrix}.$$

第三次迭代:

$$I(x_2) = \{2\},$$

$$A_1 = (-1, -5), b_1 = (-5), A_2 = \begin{bmatrix} -1 & -1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}, b_1 = \begin{bmatrix} -2 \\ 0 \\ 0 \end{bmatrix},$$

$$\nabla f(x_2) = \left(-\frac{32}{31}, -\frac{160}{31}\right)^{\mathrm{T}}.$$

求解辅助线性规划问题

$$\begin{cases} \min & -\frac{32}{31}d_1 - \frac{160}{31}d_2; \\ \text{s. t.} & -d_1 - 5d_2 \geqslant 0, \\ & -1 \leqslant d_1 \leqslant 1, \\ & -1 \leqslant d_2 \leqslant 1, \end{cases}$$

得最优解 $d_2 = \left(1, -\frac{1}{5}\right)^{\mathrm{T}}$.

因 $|\nabla f(x_2)^{\mathrm{T}} d_2| = 0$, 故 $x_2 = \left(\frac{35}{31}, \frac{24}{31}\right)^{\mathrm{T}}$ 为 K - T 点. 又因为问题(9.1.19)为凸规划,所以 $\bar{x} = x_2 = \left(\frac{35}{31}, \frac{24}{31}\right)^{\mathrm{T}}$ 为该问题的最优解. □

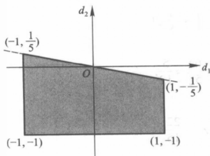


图 9.1.3 第三次迭代的图解法

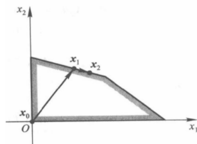


图 9.1.4 Zoutendijk 法迭代进程

9.1.2 非线性约束情形

考虑不等式约束问题

$$\begin{cases}\min & f(\boldsymbol{x});\\ \text{s.t.} & g_i(\boldsymbol{x})\geqslant 0, i=1,2,\cdots,m,\end{cases}\tag{9.1.20}$$

其中 $\boldsymbol{x}\in\mathbb{R}^n$, $f(\boldsymbol{x})$, $g_i(\boldsymbol{x})(i=1,2,\cdots,m)$均为可微函数.

设 $\boldsymbol{x}$ 是问题(9.1.20)的可行解,点 $\boldsymbol{x}$ 处起作用约束的指标集记作 $I(\boldsymbol{x})=\{i\mid g_i(\boldsymbol{x})=0, i=1,2,\cdots,m\}$. 根据定理 1.2.3 及定理 3.3.2 的证明,如果 $\boldsymbol{d}\in\mathbb{R}^n$,使

$$\begin{cases}\nabla f(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{d}<0,\\ \nabla g_i(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{d}>0, i\in I(\boldsymbol{x}),\end{cases}\tag{9.1.21}$$

则 $\boldsymbol{d}$ 是可行下降方向. 这样,求解不等式组(9.1.21)可归结为求解下列线性规划问题

$$\begin{cases}\min & z;\\ \text{s.t.} & \nabla f(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{d}-z\leqslant 0,\\ & \nabla g_i(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{d}+z\geqslant 0, i\in I(\boldsymbol{x}),\\ & -1\leqslant d_j\leqslant 1, j=1,2,\cdots,n.\end{cases}\tag{9.1.22}$$

设问题(9.1.22)的最优解为$(\bar{z},\bar{\boldsymbol{d}})$,有下述两种情况:

(1) 若 $\bar{z}<0$,则$\bar{\boldsymbol{d}}$为 $\boldsymbol{x}$ 处的可行下降方向;

(2) 若 $\bar{z}=0$,则下面的定理将证明 $\boldsymbol{x}$ 是 Fritz John 点.

定理 9.1.3 设 $\boldsymbol{x}$ 是问题(9.1.20)的可行解,函数$f(\boldsymbol{x})$, $g_i(\boldsymbol{x})(i\in I(\boldsymbol{x}))$在 $\boldsymbol{x}$ 处可微,函数 $g_i(\boldsymbol{x})(i\notin I(\boldsymbol{x}))$在点 $\boldsymbol{x}$ 处连续,则 $\boldsymbol{x}$ 是 Fritz John 点的充要条件是问题(9.1.22)的目标函数最优值为 0.

证明 对于问题(9.1.22),目标函数最优值为 0 的充要条件是不等式组

$$\begin{cases}\nabla f(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{d}<0,\\ \nabla g_i(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{d}>0, i\in I(\boldsymbol{x})\end{cases}$$

无解,即不等式组

$$\begin{cases}\nabla f(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{d}<0,\\ -\nabla g_i(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{d}<0, i\in I(\boldsymbol{x})\end{cases}\tag{9.1.23}$$

无解.

根据 Gordan 定理,(9.1.23)式无解的充要条件是:存在不全为 0 的数 $u_0\geqslant 0$ 和 $u_i\geqslant 0, i\in I(\boldsymbol{x})$,使得

$$u_0\nabla f(\boldsymbol{x})-\sum_{i\in I(\boldsymbol{x})}u_i\nabla g_i(\boldsymbol{x})=\boldsymbol{0},$$

即 $\boldsymbol{x}$ 是 Fritz John 点. □

为确定步长 λ_k，需要求解如下带约束的一维最优化问题

$$\min_{0\leqslant\lambda\leqslant\lambda_{\max}} f(\boldsymbol{x}_k+\lambda\boldsymbol{d}_k),$$

其中

$$\lambda_{\max}=\sup\{\lambda\mid g_i(\boldsymbol{x}_k+\lambda\boldsymbol{d}_k)\geqslant 0, i=1,2,\cdots,m\}. \tag{9.1.24}$$

算法 9-2(非线性约束 Zoutendijk 法)

Step 1　选取初始数据. 给定初始可行点 $\boldsymbol{x}_0$，允许误差 $\varepsilon>0$，令 $k=0$.

Step 2　确定起作用约束. 确定点 $\boldsymbol{x}_k$ 处起作用约束的指标集 $I(\boldsymbol{x}_k)=\{i\mid g_i(\boldsymbol{x}_k)=0, i=1,2,\cdots,m\}$.

Step 3　求解辅助线性规划问题. 求解线性规划问题

$$\begin{cases}\min & z;\\ \text{s.t.} & \nabla f(\boldsymbol{x}_k)^{\mathrm T}\boldsymbol{d}-z\leqslant 0,\\ & \nabla g_i(\boldsymbol{x}_k)^{\mathrm T}\boldsymbol{d}+z\geqslant 0, i\in I(\boldsymbol{x}_k),\\ & -1\leqslant d_j\leqslant 1, j=1,2,\cdots,n,\end{cases}$$

得最优解 $(z_k,\boldsymbol{d}_k)$.

Step 4　构造可行下降方向. 若 $|z_k|\leqslant\varepsilon$，停止迭代，输出 $\boldsymbol{x}_k$；否则，得到可行下降方向 $\boldsymbol{d}_k$，进行 Step 5.

Step 5　进行一维搜索. 求解

$$\min_{0\leqslant\lambda\leqslant\lambda_{\max}} f(\boldsymbol{x}_k+\lambda\boldsymbol{d}_k),$$

得最优解 λ_k，其中 $\lambda_{\max}$ 按(9.1.24)式计算. 令 $\boldsymbol{x}_{k+1}=\boldsymbol{x}_k+\lambda_k\boldsymbol{d}_k$，$k:=k+1$，返回 Step 2.

例 9.1.2　用 Zoutendijk 法求解

$$\begin{cases}\min & f(\boldsymbol{x})=2x_1^2+2x_2^2-2x_1x_2-4x_1-6x_2;\\ \text{s.t.} & g_1(\boldsymbol{x})=-2x_1^2+x_2\geqslant 0,\\ & g_2(\boldsymbol{x})=-x_1-5x_2+5\geqslant 0,\\ & g_3(\boldsymbol{x})=x_1\geqslant 0,\\ & g_4(\boldsymbol{x})=x_2\geqslant 0,\end{cases} \tag{9.1.25}$$

取初始点 $\boldsymbol{x}_0=(0,0.75)^{\mathrm T}$，迭代三次.

解　$\nabla f(\boldsymbol{x})=\begin{bmatrix}4x_1-2x_2-4\\ 4x_2-2x_1-6\end{bmatrix}$.

第一次迭代：

$$I(\boldsymbol{x}_0)=\{3\}, \nabla f(\boldsymbol{x}_0)=(-5.5,-3)^{\mathrm T}, \nabla g_3(\boldsymbol{x}_0)=(1,0)^{\mathrm T}.$$

求解辅助线性规划问题

$$\begin{cases}\min & z;\\ \text{s.t.} & -5.5d_1-3d_2\leqslant z,\\ & -d_1\leqslant z,\\ & -1\leqslant d_1\leqslant 1,\\ & -1\leqslant d_2\leqslant 1,\end{cases}$$

得到最优解 $\boldsymbol{d}_0=(1,-1)^{\mathrm{T}}, z_0=-1$.

把 $\boldsymbol{x}_0+\lambda\boldsymbol{d}_0=(\lambda,0.75-\lambda)^{\mathrm{T}}$ 代入问题(9.1.25)的约束条件,有

$$\begin{cases}-2\lambda^2+(0.75-\lambda)\geqslant 0,\\ -\lambda-5(0.75-\lambda)+5\geqslant 0,\\ \lambda\geqslant 0,\\ 0.75-\lambda\geqslant 0,\end{cases}$$

得到使 $\boldsymbol{x}_0+\lambda\boldsymbol{d}_0$ 为问题(9.1.25)的可行解的 λ 的最大值

$$\lambda_{\max}=\frac{-1+\sqrt{7}}{4}\approx 0.411\ 4.$$

求解

$$\min_{0\leqslant\lambda\leqslant 0.411\ 4} f(\boldsymbol{x}_0+\lambda\boldsymbol{d}_0)=6\lambda^2-2.5\lambda-3.375,$$

得 $\lambda_0=0.208\ 3$. 所以

$$\boldsymbol{x}_1=\boldsymbol{x}_0+\lambda_0\boldsymbol{d}_0=(0.208\ 3,0.541\ 7)^{\mathrm{T}}.$$

第二次迭代:

$$I(\boldsymbol{x}_1)=\varnothing, \nabla f(\boldsymbol{x}_1)=(-4.25,-4.25)^{\mathrm{T}}.$$

因此取 $\boldsymbol{d}_1$ 与 $\nabla f(\boldsymbol{x}_1)$ 同向,即 $\boldsymbol{d}_1=(-1,-1)$,得到 $z_1=-8.5$.

把 $\boldsymbol{x}_1+\lambda\boldsymbol{d}_1=(\lambda+0.208\ 3,\lambda+0.541\ 7)^{\mathrm{T}}$ 代入问题(9.1.25)的约束条件,求得

$$\lambda_{\max}=0.347\ 2.$$

求解

$$\min_{0\leqslant\lambda\leqslant 0.347\ 2} f(\boldsymbol{x}_0+\lambda\boldsymbol{d}_1)=2\lambda^2-8.5\lambda-3.635\ 4.$$

得 $\lambda_1=0.347\ 2$. 所以

$$\boldsymbol{x}_2=\boldsymbol{x}_1+\lambda_1\boldsymbol{d}_1=(0.555\ 5,0.888\ 9)^{\mathrm{T}}.$$

第三次迭代:

$I(\boldsymbol{x}_2)=\{2\}, \nabla f(\boldsymbol{x}_2)=(-3.555\ 8,-3.555\ 4)^{\mathrm{T}}, \nabla g_2(\boldsymbol{x}_2)=(-1,-5)^{\mathrm{T}}$,求解辅助线性规划问题

$$\begin{cases}\min & z;\\ \text{s.t.} & -3.5558d_1-3.5554\leqslant z,\\ & -d_1-5d_2\leqslant z,\\ & -1\leqslant d_1\leqslant 1,\\ & -1\leqslant d_2\leqslant 1,\end{cases}$$

得最优解

$$\boldsymbol{d}_2=(1,-0.5325)^{\mathrm{T}},z_2=-1.6625.$$

类似地容易求出

$$\lambda_{\max}=0.0925.$$

再求解

$$\min_{0\leqslant\lambda\leqslant 0.0925} f(\boldsymbol{x}_2+\lambda\boldsymbol{d}_2)=3.6321\lambda^2-1.6625\lambda-6.3455,$$

得 $\lambda_2=0.0925$. 所以

$$\boldsymbol{x}_3=\boldsymbol{x}_2+\lambda_2\boldsymbol{d}_2=(0.6480,0.8396)^{\mathrm{T}}.$$ □

计算实践和理论分析表明,该算法可能失效或出现锯齿现象,使算法收敛很慢甚至不收敛到最优点或 K－T 点.

9.1.3 Zoutendijk 法的改进

对于线性和非线性不等式约束问题,前面我们仅使用起作用约束来确定搜索方向. 当某迭代点在一个约束的边界上时,如果可行方向取得不恰当,那么沿该方向可能因接近另一个约束边界而只能作一个微小的移动,否则,就会使迭代点跑出边界. 为防止这一现象发生,设想在约束条件的边界上设立一道"安全带",迭代点进入"安全带"时,只允许它往可行域内部移动,而不许向边界靠近. 为此引入 ε 起作用约束的概念,即在构造可行方向时,既把通过当前迭代点的约束边界看作起作用约束,也把充分靠近当前迭代点的边界约束考虑在内.

对于给定的 $\varepsilon_k>0$,称

$$I_{\varepsilon_k}(\boldsymbol{x}_k)=\{i\mid 0\leqslant g_i(\boldsymbol{x}_k)\leqslant\varepsilon_k\}$$

为点 $\boldsymbol{x}_k$ 处的 ε_k 起作用约束指标集.

算法 9－3(ε 起作用约束可行方向法)

Step 1　给定初始可行点 $\boldsymbol{x}_0$,允许误差 $\varepsilon_0>0,\varepsilon>0,\varepsilon'>0$,令 $k=0$.

Step 2　确定 ε_k 起作用约束指标集

$$I_{\varepsilon_k}(\boldsymbol{x}_k)=\{i\mid 0\leqslant g_i(\boldsymbol{x}_k)\leqslant\varepsilon_k\}.$$

计算 $\nabla f(\boldsymbol{x}_k)$.

Step 3　求解辅助线性规划问题

$$
\begin{cases}
\min & z; \\
\text{s. t.} & \nabla f(\boldsymbol{x}_k)^{\mathrm{T}}\boldsymbol{d} - z \leqslant 0, \\
& \nabla g_i(\boldsymbol{x}_k)^{\mathrm{T}}\boldsymbol{d} + \eta_i z \geqslant 0, i \in I_{\varepsilon_k}(\boldsymbol{x}_k), \\
& -1 \leqslant d_j \leqslant 1, j = 1,2,\cdots,n,
\end{cases}
$$

得最优解$(z_k, \boldsymbol{d}_k)$. 其中,当 ε 起作用约束为线性约束时,取 $\eta_i = 0$;当 ε 起作用约束为非线性约束时,取 $\eta_i = 1$.

Step 4　若 $|z_k| < \varepsilon$,转 Step 6;否则,进行 Step 5.

Step 5　求解

$$
\min_{0 \leqslant \lambda \leqslant \lambda_{\max}} f(\boldsymbol{x}_k + \lambda \boldsymbol{d}_k),
$$

得最优解 λ_k,其中 $\lambda_{\max}$ 按(9.1.24)式计算. 令 $\boldsymbol{x}_{k+1} = \boldsymbol{x}_k + \lambda_k \boldsymbol{d}_k$, $k := k+1$,返回 Step 2.

Step 6　若 $\varepsilon_k \leqslant \varepsilon'$,停止迭代,输出 $\boldsymbol{x}_k$;否则,令 $\varepsilon_{k+1} = \alpha\varepsilon_k (0 < \alpha < 1)$, $k := k+1$,返回 Step 2.

为防止锯齿现象,还可考虑起作用约束和不起作用约束在确定搜索方向中都起作用. 这种全作用约束方向法是 Topkis 和 Veinott(1967)提出并保证收敛于 Fritz John 点. 对于问题(9.1.20)的一个可行点 $\boldsymbol{x}$,通过求解下面修改的辅助线性规划问题来确定搜索方向:

$$
\begin{cases}
\min & z; \\
\text{s. t.} & \nabla f(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{d} - z \leqslant 0, \\
& \nabla g_i(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{d} + z \geqslant -g_i(\boldsymbol{x}), i = 1,2,\cdots,m \\
& -1 \leqslant d_j \leqslant 1, j = 1,2,\cdots,n.
\end{cases}
\tag{9.1.26}
$$

定理 9.1.4　设 $\boldsymbol{x}$ 是问题(9.1.20)的可行解,函数 $f(\boldsymbol{x})$, $g_i(\boldsymbol{x})$ $(i = 1,2,\cdots,m)$ 在点 $\boldsymbol{x}$ 处可微,$(\bar{z}, \bar{\boldsymbol{d}})$ 是问题(9.1.26)的最优解. 如果 $\bar{z} < 0$,则 $\bar{\boldsymbol{d}}$ 为 $\boldsymbol{x}$ 处的可行下降方向;而且,$\bar{z} = 0$ 当且仅当 $\boldsymbol{x}$ 是一个 Fritz John 点.

证明　记 $I(\boldsymbol{x}) = \{i \mid g_i(\boldsymbol{x}) = 0, i = 1,2,\cdots,m\}$. 若 $\bar{z} < 0$,则对一切 $i \in I(\boldsymbol{x})$ 有 $\nabla g_i(\boldsymbol{x})^{\mathrm{T}}\bar{\boldsymbol{d}} > 0$,又 $\nabla f(\boldsymbol{x})^{\mathrm{T}}\bar{\boldsymbol{d}} < 0$,故(9.1.21)式成立,于是,$\bar{\boldsymbol{d}}$ 为可行下降方向.

我们知道,对一切 $i \in I(\boldsymbol{x})$, $g_i(\boldsymbol{x}) = 0$;对一切 $i \notin I(\boldsymbol{x})$, $g_i(\boldsymbol{x}) > 0$. 容易验证,$\bar{z} = 0$ 当且仅当不等式组

$$
\begin{cases}
\nabla f(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{d} < 0, \\
\nabla g_i(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{d} > 0, i \in I(\boldsymbol{x})
\end{cases}
$$

无解. 类似于定理 9.1.3 的证明,由 Gordan 定理知,该组无解当且仅当 $\boldsymbol{x}$ 是一个 Fritz John 点. □

算法 9－4（Topkis－Veinott 可行方向法）

Step 1　给定初始可行点 $\boldsymbol{x}_0$，允许误差 $\varepsilon>0$，令 $k=0$.

Step 2　求解辅助线性规划问题

$$\begin{cases}\min & z;\\ \text{s.t.} & \nabla f(\boldsymbol{x}_k)^{\mathrm{T}}\boldsymbol{d}-z\leqslant 0,\\ & \nabla g_i(\boldsymbol{x}_k)^{\mathrm{T}}\boldsymbol{d}+z\geqslant -g_i(\boldsymbol{x}_k),i=1,2,\cdots,m,\\ & -1\leqslant d_j\leqslant 1,j=1,2,\cdots,n,\end{cases}$$

得最优解 $(z_k,\boldsymbol{d}_k)$.

Step 3　若 $|z_k|\leqslant\varepsilon$，则停止计算，输出 $\boldsymbol{x}_k$；否则，进行 Step 4.

Step 4　求解

$$\min_{0\leqslant\lambda\leqslant\lambda_{\max}} f(\boldsymbol{x}_k+\lambda\boldsymbol{d}_k),$$

得最优解 λ_k，其中 $\lambda_{\max}$ 按(9.1.24)式计算. 令 $\boldsymbol{x}_{k+1}=\boldsymbol{x}_k+\lambda_k\boldsymbol{d}_k$，$k:=k+1$，转Step 2.

9.2　梯度投影法

考虑线性约束问题

$$\begin{cases}\min & f(\boldsymbol{x});\\ \text{s.t.} & \boldsymbol{Ax}\geqslant\boldsymbol{b},\\ & \boldsymbol{Ex}=\boldsymbol{e},\end{cases}\tag{9.2.1}$$

其中 $f(\boldsymbol{x})$ 为可微函数，$\boldsymbol{A}$ 为 $m\times n$ 矩阵，$\boldsymbol{E}$ 为 $l\times n$ 矩阵，$\boldsymbol{x}\in\mathbb{R}^n$，$\boldsymbol{b}\in\mathbb{R}^m$，$\boldsymbol{e}\in\mathbb{R}^l$.

我们知道，目标函数的最速下降方向是负梯度方向. 但是，在有约束情况下，沿最速下降方向移动可能导致非可行点. Rosen(1960)提出，并由 Goldfarb 和 Lapidus(1968)加以改进的梯度投影法(gradient projection method)是对负梯度通过既改进目标函数值同时又保持可行性的方式进行投影. 具体说来就是，当迭代点在可行域内部时，取该点处的负梯度方向为可行下降方向；当迭代点在可行域边界上时，取该点处负梯度方向在可行域边界上的投影产生一个可行下降方向(见图 9.2.1).

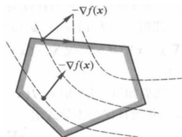


图 9.2.1　投影梯度法的搜索方向

怎样得到一个方向在某子空间上的投影方向呢? 下面我们引入投影矩阵(projection matrix)的概念.

9.2.1 投影矩阵

设 V 是$\mathbb{R}^n$ 的子空间，对于任意 $\boldsymbol{x}\in\mathbb{R}^n$，可惟一分解为

$$\boldsymbol{x}=\boldsymbol{p}+\boldsymbol{q},\boldsymbol{p}\in V,\boldsymbol{q}\in V^{\perp}.$$

若矩阵 $\boldsymbol{P}\in\mathbb{R}^{n\times n}$，使 $\boldsymbol{p}=\boldsymbol{Px}$，则称 $\boldsymbol{P}$ 为$\mathbb{R}^n$ 到子空间 V 的正交投影矩阵，简称为投影矩阵(projection matrix)，$\boldsymbol{p}$ 称为 $\boldsymbol{x}$ 在 V 上的投影.

设 $\boldsymbol{M}$ 是 $m\times n(m\leqslant n)$ 行满秩矩阵，$\boldsymbol{M}^{\mathrm{T}}$ 的 m 个列向量生成的子空间记作 V_M，V_M 的正交补子空间用 $V_M^{\perp}$ 表示，则

$$V_M^{\perp}=\{\boldsymbol{x}\mid \boldsymbol{Mx}=\boldsymbol{0}\}$$

是 $\boldsymbol{M}$ 的零空间. 对于任意 $\boldsymbol{x}\in\mathbb{R}^n$，可惟一分解为

$$\boldsymbol{x}=\boldsymbol{p}+\boldsymbol{q},\boldsymbol{p}\in V_M,\boldsymbol{q}\in V_M^{\perp},\tag{9.2.2}$$

$\boldsymbol{p},\boldsymbol{q}$ 分别是 $\boldsymbol{x}$ 在 V_M 和 $V_M^{\perp}$ 上的投影，记

$$\boldsymbol{p}=\boldsymbol{Px},\boldsymbol{q}=\boldsymbol{Qx},\tag{9.2.3}$$

则 $\boldsymbol{P},\boldsymbol{Q}$ 分别是$\mathbb{R}^n$ 到 V_M 和$\mathbb{R}^n$ 到 $V_M^{\perp}$ 上的投影矩阵.

因 $\boldsymbol{p}$ 可表示为 $\boldsymbol{M}^{\mathrm{T}}$ 的 m 个列向量的线性组合，即存在系数向量 $\boldsymbol{\eta}\in\mathbb{R}^m$，使

$$\boldsymbol{p}=\boldsymbol{M}^{\mathrm{T}}\boldsymbol{\eta}.\tag{9.2.4}$$

将(9.2.4)式代入(9.2.2)式，然后两端左乘 $\boldsymbol{M}$，并注意到 $\boldsymbol{Mq}=\boldsymbol{0}$，则

$$\boldsymbol{Mx}=\boldsymbol{MM}^{\mathrm{T}}\boldsymbol{\eta}.$$

因 $\boldsymbol{M}$ 行满秩，有

$$\boldsymbol{\eta}=(\boldsymbol{MM}^{\mathrm{T}})^{-1}\boldsymbol{Mx}.$$

将其代入(9.2.4)式得

$$\boldsymbol{p}=\boldsymbol{M}^{\mathrm{T}}(\boldsymbol{MM}^{\mathrm{T}})^{-1}\boldsymbol{Mx},$$

将上式代入(9.2.2)式得

$$\boldsymbol{q}=[\boldsymbol{I}-\boldsymbol{M}^{\mathrm{T}}(\boldsymbol{MM}^{\mathrm{T}})^{-1}\boldsymbol{M}]\boldsymbol{x},$$

再根据(9.2.3)式得

$$\boldsymbol{P}=\boldsymbol{M}^{\mathrm{T}}(\boldsymbol{MM}^{\mathrm{T}})^{-1}\boldsymbol{M},$$
$$\boldsymbol{Q}=\boldsymbol{I}-\boldsymbol{M}^{\mathrm{T}}(\boldsymbol{MM}^{\mathrm{T}})^{-1}\boldsymbol{M}.$$

容易验证，投影矩阵 $\boldsymbol{P},\boldsymbol{Q}$ 是幂等的对称矩阵. 反之，一个幂等的对称矩阵必是一个投影矩阵，这就是下面的定理.

定理 9.2.1 设 $\boldsymbol{P}$ 为 n 阶矩阵，$\boldsymbol{P}^{\mathrm{T}}=\boldsymbol{P}$，且 $\boldsymbol{P}^2=\boldsymbol{P}$，则 $\boldsymbol{P}$ 为投影矩阵.

证明 记

$$V_1=\{\boldsymbol{p}\mid\boldsymbol{p}=\boldsymbol{Px},\boldsymbol{x}\in\mathbb{R}^n\},V_2=\{\boldsymbol{q}\mid\boldsymbol{q}=\boldsymbol{Qx},\boldsymbol{x}\in\mathbb{R}^n\},$$

其中，$\boldsymbol{Q}=\boldsymbol{I}-\boldsymbol{P}$，则 V_1,V_2 是$\mathbb{R}^n$ 的子空间，且对任意 $\boldsymbol{x}\in\mathbb{R}^n$，可分解为

$$\boldsymbol{x}=\boldsymbol{p}+\boldsymbol{q},\boldsymbol{p}\in V_1,\boldsymbol{q}\in V_2.$$

又对一切 $\boldsymbol{p}\in V_1,\boldsymbol{q}\in V_2$，有

$$\boldsymbol{p}^{\mathrm{T}}\boldsymbol{q}=\boldsymbol{x}^{\mathrm{T}}\boldsymbol{P}^{\mathrm{T}}\boldsymbol{Q}\boldsymbol{x}=\boldsymbol{x}^{\mathrm{T}}\boldsymbol{P}(\boldsymbol{I}-\boldsymbol{P})\boldsymbol{x}=\boldsymbol{x}^{\mathrm{T}}(\boldsymbol{P}-\boldsymbol{P}^2)\boldsymbol{x}=0,$$

故 $V_2=V_1^{\perp}$，从而 $\boldsymbol{P}$ 为$\mathbb{R}^n$ 到 V_1 的投影矩阵. □

投影矩阵具有下述性质：

(1) 若 $\boldsymbol{P}$ 为投影矩阵，则 $\boldsymbol{P}$ 为半正定矩阵；

(2) $\boldsymbol{P}$ 为投影矩阵的充要条件是 $\boldsymbol{Q}=\boldsymbol{I}-\boldsymbol{P}$ 为投影矩阵.

9.2.2 梯度投影法

定理 9.2.2 设 $\boldsymbol{x}$ 是问题(9.2.1)的可行解，在点 $\boldsymbol{x}$ 处有 $\boldsymbol{A}_1\boldsymbol{x}=\boldsymbol{b}_1,\boldsymbol{A}_2\boldsymbol{x}>\boldsymbol{b}_2$，其中

$$\boldsymbol{A}=\begin{bmatrix}\boldsymbol{A}_1\\\boldsymbol{A}_2\end{bmatrix},\boldsymbol{b}=\begin{bmatrix}\boldsymbol{b}_1\\\boldsymbol{b}_2\end{bmatrix}.$$

又设 $\boldsymbol{M}=\begin{bmatrix}\boldsymbol{A}_1\\\boldsymbol{E}\end{bmatrix}$为行满秩矩阵，记投影矩阵

$$\boldsymbol{P}=\boldsymbol{I}-\boldsymbol{M}^{\mathrm{T}}(\boldsymbol{M}\boldsymbol{M}^{\mathrm{T}})^{-1}\boldsymbol{M},$$

若 $\boldsymbol{P}\nabla f(\boldsymbol{x})\neq\boldsymbol{0}$，则 $\boldsymbol{d}=-\boldsymbol{P}\nabla f(\boldsymbol{x})$是点 $\boldsymbol{x}$ 处的可行下降方向.

证明 因 $\boldsymbol{P}$ 为投影矩阵，$\boldsymbol{P}\nabla f(\boldsymbol{x})\neq\boldsymbol{0}$，故

$$\nabla f(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{d}=-\nabla f(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{P}\nabla f(\boldsymbol{x})=-\|\boldsymbol{P}\nabla f(\boldsymbol{x})\|^2<0,$$

即 $\boldsymbol{d}$ 为下降方向.

根据假设，又有

$$\begin{aligned}\boldsymbol{M}\boldsymbol{d}&=-\boldsymbol{M}\boldsymbol{P}\nabla f(\boldsymbol{x})=-\boldsymbol{M}(\boldsymbol{I}-\boldsymbol{M}^{\mathrm{T}}(\boldsymbol{M}\boldsymbol{M}^{\mathrm{T}})^{-1}\boldsymbol{M})\nabla f(\boldsymbol{x})\\&=(-\boldsymbol{M}+\boldsymbol{M})\nabla f(\boldsymbol{x})=\boldsymbol{0},\end{aligned}$$

即 $\boldsymbol{A}_1\boldsymbol{d}=\boldsymbol{0},\boldsymbol{E}\boldsymbol{d}=\boldsymbol{0}$，根据定理 9.1.1，$\boldsymbol{d}$ 是点 $\boldsymbol{x}$ 处的可行方向. 所以，$\boldsymbol{d}$ 是可行下降方向. □

上述定理表明，在 $\boldsymbol{P}\nabla f(\boldsymbol{x})\neq\boldsymbol{0}$ 的假设下，如果 $\boldsymbol{M}$ 为空，即迭代点在可行域内部时，负梯度方向 $\boldsymbol{d}=-\nabla f(\boldsymbol{x})$为可行下降方向；如果 $\boldsymbol{M}$ 非空，即迭代点在某些约束的边界上时，该点处的负梯度在 $\boldsymbol{M}$ 的零空间上的投影为可行下降方向.

如果 $\boldsymbol{P}\nabla f(\boldsymbol{x})=\boldsymbol{0}$ 时，则由下面的定理知，或者 $\boldsymbol{x}$ 是 K－T 点，或者可以构造新的投影矩阵以便求得可行下降方向.

定理 9.2.3 设 $\boldsymbol{x}$ 是问题(9.2.1)的一个可行解，在点 $\boldsymbol{x}$ 处，有 $\boldsymbol{A}_1\boldsymbol{x}=\boldsymbol{b}_1,\boldsymbol{A}_2\boldsymbol{x}>\boldsymbol{b}_2$，其中

$$\boldsymbol{A}=\begin{bmatrix}\boldsymbol{A}_1\\\boldsymbol{A}_2\end{bmatrix},\boldsymbol{b}=\begin{bmatrix}\boldsymbol{b}_1\\\boldsymbol{b}_2\end{bmatrix}.$$

又设 $\boldsymbol{M}=\begin{bmatrix}\boldsymbol{A}_1\\\boldsymbol{E}\end{bmatrix}$为行满秩矩阵，令

$$P = I - M^{T}(MM^{T})^{-1}M, W = (MM^{T})^{-1}M\nabla f(x) = \begin{bmatrix} u \\ v \end{bmatrix},$$

其中 u 和 v 分别对应于矩阵 A_1 和 E,设 $P\nabla f(x) = 0$,则

(1) 如果 $u \geqslant 0$,那么 x 为 K-T 点.

(2) 如果 u 中含有负分量,不妨设 $u_j < 0$,这时从 A_1 去掉 u_j 对应的行,得到 $\hat{A}_1$,令

$$\hat{M} = \begin{bmatrix} \hat{A}_1 \\ E \end{bmatrix}, \hat{P} = I - \hat{M}^{T}(\hat{M}\hat{M}^{T})^{-1}\hat{M}, d = -\hat{P}\nabla f(x),$$

那么 d 为点 x 处的可行下降方向.

证明 先证(1). 由于 $P\nabla f(x) = 0$,因此有

$$\begin{aligned} 0 &= P\nabla f(x) = [I - M^{T}(MM^{T})^{-1}M]\nabla f(x) \\ &= \nabla f(x) - M^{T}(MM^{T})^{-1}M\nabla f(x) \\ &= \nabla f(x) - (A_1^{T}, E^{T})\begin{bmatrix} u \\ v \end{bmatrix} = \nabla f(x) - A_1^{T}u - E^{T}v, \end{aligned} \tag{9.2.5}$$

当 $u \geqslant 0$ 时,(9.2.5)式恰为 K-T 条件,因此 x 是 K-T 点.

再证(2). 设 $u_j < 0$,先证 $\hat{P}\nabla f(x) \neq 0$. 用反证法. 假设 $\hat{P}\nabla f(x) = 0$,由 $\hat{P}$ 的定义可以推出

$$\begin{aligned} 0 = \hat{P}\nabla f(x) &= (I - \hat{M}^{T}(\hat{M}\hat{M}^{T})^{-1}\hat{M})\nabla f(x) \\ &= \nabla f(x) - \hat{M}^{T}\hat{W}, \end{aligned} \tag{9.2.6}$$

其中 $\hat{W} = (\hat{M}\hat{M}^{T})^{-1}\hat{M}\nabla f(x)$. 设 r_j^{T} 是 A_1 中对应 u_j 的行向量(第 j 行),由于

$$A_1^{T}u + E^{T}v = \hat{A}_1^{T}\bar{u} + u_j r_j + E^{T}v = \hat{M}^{T}\overline{W} + u_j r_j, \tag{9.2.7}$$

其中 $\overline{W} = \begin{bmatrix} \bar{u} \\ v \end{bmatrix}$,$\bar{u}$ 为 u 中划去分量 u_j 的向量. 因此,将(9.2.7)式代入(9.2.5)式得

$$0 = \nabla f(x) - \hat{M}^{T}\overline{W} - u_j r_j, \tag{9.2.8}$$

将(9.2.6)式与(9.2.8)式相减得

$$0 = \hat{M}^{T}(\overline{W} - \hat{W}) + u_j r_j.$$

上式右端为 M^{T} 的列向量的线性组合,且至少有一个系数 $u_j \neq 0 (u_j < 0)$,由此得出 M^{T} 的列向量组线性相关,这与 M 为行满秩矩阵的假设矛盾. 因此,必有 $\hat{P}\nabla f(x) \neq 0$.

由于 $\hat{P}$ 为投影矩阵,$\hat{P}\nabla f(x) \neq 0$,因此 $d = -\hat{P}\nabla f(x)$ 是下降方向. 再证 d 为

可行方向. 因为

$$\hat{M}d = -\hat{M}\hat{P}\nabla f(x) = -\hat{M}(I-\hat{M}(\hat{M}\hat{M}^{T})^{-1}\hat{M})\nabla f(x)$$
$$= -(\hat{M}-\hat{M})\nabla f(x) = 0,$$

所以

$$\hat{A}_1 d = 0, Ed = 0. \tag{9.2.9}$$

又由(9.2.8)式两端左乘 $r_j^T\hat{P}$,得

$$r_j^T\hat{P}\nabla f(x) - r_j^T\hat{P}\hat{M}^T\overline{W} - u_j r_j^T\hat{P}r_j = 0.$$

因$\hat{P}\hat{M}^T=0$, $d=-\hat{P}\nabla f(x)$,上式即

$$r_j^T d + u_j r_j^T\hat{P}r_j = 0. \tag{9.2.10}$$

因$\hat{P}$半正定,故 $r_j^T\hat{P}r_j \geqslant 0$,又 $u_j<0$,则由(9.2.10)式知

$$r_j^T d \geqslant 0. \tag{9.2.11}$$

由(9.2.9)和(9.2.11)两式得

$$A_1 d \geqslant 0, Ed = 0.$$

根据定理9.1.1,d 为可行方向. 所以,d 为 x 处的可行下降方向. □

上述两个定理给出了梯度投影法中可行下降方向的构造以及算法停止于K-T点的条件. 梯度投影法中沿可行下降方向的一维搜索与Zoutendijk法中的一维搜索相同.

算法9-5(梯度投影法)

Step 1　选取初始数据. 选取初始可行点 x_0,允许误差 $\varepsilon>0$,令 $k=0$.

Step 2　分解出起作用约束. 在点 x_k 处,将 A 和 b 分解为

$$A=\begin{bmatrix}A_1\\A_2\end{bmatrix}, b=\begin{bmatrix}b_1\\b_2\end{bmatrix},$$

使得 $A_1x_k=b_1, A_2x_k>b_2$.

Step 3　确定投影矩阵. 令

$$M=\begin{bmatrix}A_1\\E\end{bmatrix},$$

若 M 为空,则令 $P=I_n$(n 阶单位矩阵),否则,令

$$P=I_n-M^T(MM^T)^{-1}M.$$

Step 4　构造可行下降方向. 令 $d_k=-P\nabla f(x_k)$,若 $\|d_k\|\leqslant\varepsilon$,则转Step 5;否则,进行Step 6.

Step 5　检查起作用的约束数,判定迭代点 x_k 是否在可行域内部. 若 M 为空,则停止迭代,输出 x_k;否则,令

$$W=(MM^{\mathrm{T}})^{-1}M\nabla f(x_k)=\begin{bmatrix}u\\v\end{bmatrix},$$

其中 u,v 分别对应于 A_1 和 E. 如果 $u\geqslant 0$,则停止迭代,输出 x_k;如果 $u\ngeqslant 0$,则在 u 中选择一个负分量,比如 u_j,修正 A_1,即去掉 A_1 中对应 u_j 的行以后得矩阵的 $\hat{A}_1$,令 $A_1:=\hat{A}_1$,转 Step 3.

Step 6　进行一维搜索. 计算 $\hat{b}=b_2-A_2x_k,\hat{d}=A_2d_k$,求解

$$\min_{0\leqslant\lambda\leqslant\lambda_{\max}} f(x_k+\lambda d_k),$$

得最优解 λ_k,其中

$$\lambda_{\max}=\begin{cases}\min\limits_i\left\{\dfrac{\hat{b}_i}{\hat{d}_i}\,\middle|\,\hat{d}_i<0\right\}, & \text{当}\hat{d}\ngeqslant 0\text{ 时},\\ \infty, & \text{当}\hat{d}\geqslant 0\text{ 时},\end{cases}$$

令 $x_{k+1}=x_k+\lambda_k d_k,k:=k+1$,转 Step 2.

例 9.2.1　用梯度投影法求解问题(9.1.19),要求选取初始可行点 $x_0=(0,0)^{\mathrm{T}}$.

解　$\nabla f(x)=\begin{bmatrix}4x_1-2x_2-4\\-2x_1+4x_2-6\end{bmatrix}$.

第一次迭代:

在 x_0 处,起作用约束指标集 $I(x_0)=\{3,4\}$,对应有

$$A_1=\begin{bmatrix}1&0\\0&1\end{bmatrix},b_1=\begin{bmatrix}0\\0\end{bmatrix},A_2=\begin{bmatrix}-1&-1\\-1&-5\end{bmatrix},b_1=\begin{bmatrix}-2\\-5\end{bmatrix},$$

$$\nabla f(x_0)=(-4,-6)^{\mathrm{T}}.$$

投影矩阵(若 A_1 为非奇异方阵,则 $P=0$)

$$P=I-A_1^{\mathrm{T}}(A_1A_1^{\mathrm{T}})^{-1}A_1=\begin{bmatrix}0&0\\0&0\end{bmatrix}.$$

令 $d_0=-P\nabla f(x_0)=(0,0)^{\mathrm{T}}$,

$$W=(A_1,A_1^{\mathrm{T}})^{-1}A_1\nabla f(x_0)=\begin{bmatrix}-4\\-6\end{bmatrix}=\begin{bmatrix}u_1\\u_2\end{bmatrix}.$$

因 W 中含负分量,故修正 A_1. 在 A_1 中去掉 $u_2=-6$ 对应的行,即第 2 行,得$\hat{A}_1=(1,0)$. 再求投影矩阵

$$P=I-A_1^{\mathrm{T}}(A_1A_1^{\mathrm{T}})^{-1}A_1=\begin{bmatrix}1&0\\0&1\end{bmatrix}-\begin{bmatrix}1\\0\end{bmatrix}\left[(1,0)\begin{bmatrix}1\\0\end{bmatrix}\right]^{-1}(1,0)$$

$$=\begin{bmatrix}1&0\\0&1\end{bmatrix}-\begin{bmatrix}1&0\\0&0\end{bmatrix}=\begin{bmatrix}0&0\\0&1\end{bmatrix},$$

令 $\boldsymbol{d}_0 = -\boldsymbol{P}\nabla f(\boldsymbol{x}_0) = -\begin{bmatrix}0 & 0\\0 & 1\end{bmatrix}\begin{bmatrix}-4\\-6\end{bmatrix} = \begin{bmatrix}0\\6\end{bmatrix}$. 沿 $\boldsymbol{d}_0$ 进行一维搜索,计算

$$\hat{\boldsymbol{b}} = \boldsymbol{b}_2 - \boldsymbol{A}_2\boldsymbol{x}_0 = \begin{bmatrix}-2\\-5\end{bmatrix} - \begin{bmatrix}-1 & -1\\-1 & -5\end{bmatrix}\begin{bmatrix}0\\0\end{bmatrix} = \begin{bmatrix}-2\\-5\end{bmatrix},$$

$$\hat{\boldsymbol{d}} = \boldsymbol{A}_2\boldsymbol{d}_0 = \begin{bmatrix}-1 & -1\\-1 & -5\end{bmatrix}\begin{bmatrix}0\\6\end{bmatrix} = \begin{bmatrix}-6\\-30\end{bmatrix},$$

因此 $\lambda_{\max} = \min\left\{\frac{-2}{-6}, \frac{-5}{-30}\right\} = \frac{1}{6}$. 解

$$\min_{0 \leqslant \lambda \leqslant \frac{1}{6}} f(\boldsymbol{x}_0 + \lambda\boldsymbol{d}_0) = 72\lambda^2 - 36\lambda,$$

得最优步长 $\lambda_0 = \frac{1}{6}$. 令 $\boldsymbol{x}_1 = \boldsymbol{x}_0 + \lambda_0\boldsymbol{d}_0 = (0,1)^{\mathrm{T}}$.

第二次迭代:

$I(\boldsymbol{x}_1) = \{2,3\}$,对应有

$$\boldsymbol{A}_1 = \begin{bmatrix}-1 & -5\\1 & 0\end{bmatrix}, \boldsymbol{b}_1 = \begin{bmatrix}-5\\0\end{bmatrix}, \boldsymbol{A}_2 = \begin{bmatrix}-1 & -1\\0 & 1\end{bmatrix}, \boldsymbol{b}_1 = \begin{bmatrix}-2\\0\end{bmatrix},$$

$$\nabla f(\boldsymbol{x}_1) = (-6, -2)^{\mathrm{T}}.$$

投影矩阵

$$\boldsymbol{P} = \boldsymbol{I} - \boldsymbol{A}_1^{\mathrm{T}}(\boldsymbol{A}_1\boldsymbol{A}_1^{\mathrm{T}})^{-1}\boldsymbol{A}_1 = \begin{bmatrix}0 & 0\\0 & 0\end{bmatrix}.$$

令 $\boldsymbol{d}_1 = -\boldsymbol{P}\nabla f(\boldsymbol{x}_1) = (0,0)^{\mathrm{T}}$,

$$\boldsymbol{W} = (\boldsymbol{A}_1\boldsymbol{A}_1^{\mathrm{T}})^{-1}\boldsymbol{A}_1\nabla f(\boldsymbol{x}_1) = \begin{bmatrix}0 & -\frac{1}{5}\\1 & -\frac{1}{5}\end{bmatrix}\begin{bmatrix}-6\\-2\end{bmatrix} = \begin{bmatrix}\frac{2}{5}\\-\frac{28}{5}\end{bmatrix} = \begin{bmatrix}u_1\\u_2\end{bmatrix}.$$

因 $\boldsymbol{W}$ 中含有负分量,故从 $\boldsymbol{A}_1$ 划去 $u_2 = -\frac{28}{5}$ 所对应的第二行,得 $\hat{\boldsymbol{A}}_1 = (-1, -5)$. 令

$$\begin{aligned}\boldsymbol{P} &= \boldsymbol{I} - \boldsymbol{A}_1^{\mathrm{T}}(\boldsymbol{A}_1\boldsymbol{A}_1^{\mathrm{T}})^{-1}\boldsymbol{A}_1\\&= \begin{bmatrix}1 & 0\\0 & 1\end{bmatrix} - \begin{bmatrix}-1\\-5\end{bmatrix}\left[(-1,-5)\begin{bmatrix}-1\\-5\end{bmatrix}\right]^{-1}(-1,-5) = \begin{bmatrix}\frac{25}{26} & -\frac{5}{26}\\-\frac{5}{26} & \frac{1}{26}\end{bmatrix},\end{aligned}$$

$$\boldsymbol{d}_1 = -\boldsymbol{P}\nabla f(\boldsymbol{x}_1) = \frac{14}{13}(5,-1)^{\mathrm{T}}.$$

不妨去掉前面的系数,取搜索方向 $\boldsymbol{d}_1 = (5,-1)^{\mathrm{T}}$. 沿 $\boldsymbol{d}_1$ 进行一维搜索,计算

$$\hat{b} = b_2 - A_2 x_1 = \begin{bmatrix} -2 \\ 0 \end{bmatrix} - \begin{bmatrix} -1 & -1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} -1 \\ -1 \end{bmatrix},$$

$$\hat{d} = A_2 d_1 = \begin{bmatrix} -1 & -1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 5 \\ -1 \end{bmatrix} = \begin{bmatrix} -4 \\ -1 \end{bmatrix},$$

$$\lambda_{\max} = \min\left\{ \frac{-1}{-4}, \frac{-1}{-1} \right\} = \frac{1}{4}.$$

解 $\min\limits_{0 \leqslant \lambda \leqslant \frac{1}{4}} f(x_1 + \lambda d_1) = 62\lambda^2 - 28\lambda - 4$，得到 $\lambda_1 = \frac{7}{31}$. 令

$$x_2 = x_1 + \lambda_1 d_1 = \left(\frac{35}{31}, \frac{24}{31} \right)^{\mathrm{T}}.$$

第三次迭代：

在 x_2 处，$I(x_2) = \{1\}$，对应有

$$A_1 = (-1, -5), b_1 = (-5), A_2 = \begin{bmatrix} -1 & -1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}, b_1 = \begin{bmatrix} -2 \\ 0 \\ 0 \end{bmatrix},$$

$$\nabla f(x_2) = \left(-\frac{32}{31}, -\frac{160}{31} \right)^{\mathrm{T}},$$

投影矩阵

$$P = I - A_1^{\mathrm{T}} (A_1 A_1^{\mathrm{T}})^{-1} A_1$$

$$= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \begin{bmatrix} -1 \\ -5 \end{bmatrix} \left[(-1, -5) \begin{bmatrix} -1 \\ -5 \end{bmatrix} \right]^{-1} (-1, -5) = \begin{bmatrix} \frac{25}{26} & -\frac{5}{26} \\ -\frac{5}{26} & \frac{1}{26} \end{bmatrix},$$

$$d_2 = -P \nabla f(x_2) = (0,0)^{\mathrm{T}}, W = (A_1, A_1^{\mathrm{T}})^{-1} A_1 \nabla f(x_2) = \frac{32}{31} > 0.$$

由定理 9.2.3 知，$x_2 = \left(\frac{35}{31}, \frac{24}{31} \right)^{\mathrm{T}}$ 为 K－T 点. □

由于本例为凸规划，故 x_2 为全局最优解.

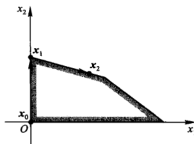


图 9.2.2 投影梯度法迭代进程

9.3 既约梯度法

1963 年，Wolfe 将线性规划的单纯形法推广到具有非线性目标函数的问题，提出了产生可行下降方向的另一类方法，称为既约梯度法（reduced gradient method）.

考虑问题

$$\begin{cases} \min & f(\boldsymbol{x}); \\ \text{s.t.} & \boldsymbol{A}\boldsymbol{x}=\boldsymbol{b}, \\ & \boldsymbol{x}\geqslant \boldsymbol{0}, \end{cases} \tag{9.3.1}$$

其中 $\boldsymbol{A}$ 为 $m\times n$ 矩阵，$m\leqslant n$，$\boldsymbol{b}\in\mathbb{R}^m$，$\boldsymbol{x}\in\mathbb{R}^n$.

9.3.1 用既约梯度构造可行下降方向

设问题(9.3.1)的约束是非退化的，即假设 $\boldsymbol{A}$ 的任意 m 个列均线性无关，并且每个基本可行解均有 m 个正分量. 这时 $\operatorname{rank}\boldsymbol{A}=m$.

设 $\boldsymbol{x}$ 为问题(9.3.1)的任一可行解，$\boldsymbol{B}$ 是 $m\times m$ 可逆矩阵（基矩阵），将矩阵 $\boldsymbol{A}$ 和向量 $\boldsymbol{x}$ 作相应划分：

$$\boldsymbol{A}=(\boldsymbol{B},\boldsymbol{N}),\quad \boldsymbol{x}=\begin{bmatrix}\boldsymbol{x}_B\\ \boldsymbol{x}_N\end{bmatrix},$$

其中 $\boldsymbol{x}_B$ 的分量称为基变量，$\boldsymbol{x}_N$ 的分量称为非基变量，并且由假设 $\boldsymbol{x}_B>\boldsymbol{0}$ 知，$\boldsymbol{x}$ 为非退化的可行解.

将问题(9.3.1)表述为

$$\begin{cases} \min & f(\boldsymbol{x}_B,\boldsymbol{x}_N); \\ \text{s.t.} & \boldsymbol{B}\boldsymbol{x}_B+\boldsymbol{N}\boldsymbol{x}_N=\boldsymbol{b}, \\ & \boldsymbol{x}_B,\boldsymbol{x}_N\geqslant \boldsymbol{0}. \end{cases} \tag{9.3.2}$$

由问题(9.3.2)中等式约束知

$$\boldsymbol{x}_B=\boldsymbol{B}^{-1}\boldsymbol{b}-\boldsymbol{B}^{-1}\boldsymbol{N}\boldsymbol{x}_N.$$

代入问题(9.3.2)中目标函数，记

$$F(\boldsymbol{x}_N)=f(\boldsymbol{x}_B,\boldsymbol{x}_N)=f(\boldsymbol{x}_B(\boldsymbol{x}_N),\boldsymbol{x}_N),$$

于是，问题(9.3.1)简化为仅在变量非负限制下的极小化问题

$$\begin{cases} \min & F(\boldsymbol{x}_N); \\ \text{s.t.} & \boldsymbol{x}_B=\boldsymbol{B}^{-1}\boldsymbol{b}-\boldsymbol{B}^{-1}\boldsymbol{N}\boldsymbol{x}_N\geqslant \boldsymbol{0}, \\ & \boldsymbol{x}_N\geqslant \boldsymbol{0}. \end{cases} \tag{9.3.3}$$

这是一个 $n-m$ 维问题，而且除变量非负约束外不带其他约束条件，因此，问题

(9.3.3)是比原来问题较低维的简单问题.

利用复合函数求导法则,可求得 $F(\boldsymbol{x}_N)$ 的梯度,即 $f(\boldsymbol{x})$ 的既约梯度(reduced gadient)

$$\boldsymbol{r}(\boldsymbol{x}_N)=\nabla F(\boldsymbol{x}_N)=\nabla_{\boldsymbol{x}_N}f(\boldsymbol{x})-(\boldsymbol{B}^{-1}\boldsymbol{N})^{\mathrm{T}}\nabla_{\boldsymbol{x}_B}f(\boldsymbol{x}).$$

显然,沿负既约梯度方向 $-\boldsymbol{r}(\boldsymbol{x}_N)$ 移动 $\boldsymbol{x}_N$,能使目标函数下降,但如何得到可行下降方向呢?

根据定理 9.1.1,非零向量 $\boldsymbol{d}$ 为 $\boldsymbol{x}$ 处的可行下降方向的充分条件是

$$\nabla f(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{d}<\boldsymbol{0}, \tag{9.3.4}$$

$$\boldsymbol{A}\boldsymbol{d}=\boldsymbol{0}, \tag{9.3.5}$$

$$d_j\geqslant 0,\text{当 } x_{N_j}=0 \text{ 时}. \tag{9.3.6}$$

若按照 $\boldsymbol{x}=\begin{bmatrix}\boldsymbol{x}_B\\ \boldsymbol{x}_N\end{bmatrix}$ 的分解方式,把 n 维向量 $\boldsymbol{d}$ 分解为 $\boldsymbol{d}=\begin{bmatrix}\boldsymbol{d}_B\\ \boldsymbol{d}_N\end{bmatrix}$,则由(9.3.5)式知 $\boldsymbol{B}\boldsymbol{d}_B+\boldsymbol{N}\boldsymbol{d}_N=\boldsymbol{0}$,故

$$\boldsymbol{d}_B=-\boldsymbol{B}^{-1}\boldsymbol{N}\boldsymbol{d}_N. \tag{9.3.7}$$

另一方面,

$$\begin{aligned}\nabla f(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{d}&=\nabla_{\boldsymbol{x}_B}f(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{d}_B+\nabla_{\boldsymbol{x}_N}f(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{d}_N\\&=\nabla_{\boldsymbol{x}_B}f(\boldsymbol{x})^{\mathrm{T}}(-\boldsymbol{B}^{-1}\boldsymbol{N}\boldsymbol{d}_N)+\nabla_{\boldsymbol{x}_N}f(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{d}_N\\&=\boldsymbol{r}(\boldsymbol{x}_N)^{\mathrm{T}}\boldsymbol{d}_N,\end{aligned}$$

由(9.3.4)式知,若 $\boldsymbol{d}$ 满足

$$\boldsymbol{r}(\boldsymbol{x}_N)^{\mathrm{T}}\boldsymbol{d}_N<0, \tag{9.3.8}$$

并同时满足(9.3.6)式和(9.3.7)式,那么 $\boldsymbol{d}$ 为可行下降方向.

如果简单地取 $\boldsymbol{d}_N=-\boldsymbol{r}(\boldsymbol{x}_N)$,则(9.3.8)式成立,$\boldsymbol{d}$ 为下降方向. 但是,当某个分量 $x_{N_j}=0$ 且 $r_j(\boldsymbol{x}_N)>0$ 时,导致(9.3.6)式不成立,破坏了可行性. 所以,Wolfe 采用如下修正:

$$d_{N_j}=\begin{cases}0, & \text{若 } x_{N_j}=0 \text{ 且 } r_j(\boldsymbol{x}_N)>0,\\ -r_j(\boldsymbol{x}_N), & \text{否则}.\end{cases}$$

但 Wolfe 后来举例说明了采用上述搜索方向时算法可能收敛到非 K-T 点. 为此,Mc Cormick 对上式作出了修正,即令

$$d_{N_j}=\begin{cases}-x_{N_j}r_j(\boldsymbol{x}_N), & \text{当 } r_j(\boldsymbol{x}_N)>0 \text{ 时},\\ -r_j(\boldsymbol{x}_N), & \text{当 } r_j(\boldsymbol{x}_N)\leqslant 0 \text{ 时}.\end{cases} \tag{9.3.9}$$

容易验证,按(9.3.9)式定义的非零向量 $\boldsymbol{d}_N$ 满足(9.3.6)和(9.3.8)式. 所以,按(9.3.7)和(9.3.9)式构造的非零向量是可行下降方向,而且我们将看到 $\boldsymbol{d}=\boldsymbol{0}$ 的充要条件是 $\boldsymbol{x}$ 为 K-T 点.

定理 9.3.1 设 $\boldsymbol{x}$ 是问题(9.3.1)的可行解,$\boldsymbol{B}$ 为基矩阵,对应有 $\boldsymbol{A}=(\boldsymbol{B},\boldsymbol{N})$,$\boldsymbol{x}=\begin{bmatrix}\boldsymbol{x}_B\\ \boldsymbol{x}_N\end{bmatrix}$,假设 $\boldsymbol{x}_B>\boldsymbol{0}$,$f:\mathbb{R}^n\to\mathbb{R}$ 在点 $\boldsymbol{x}$ 处可微,又设 $\boldsymbol{d}$ 是按(9.3.7)和(9.3.9)式定义的方向,则

(1) 当 $\boldsymbol{d}\neq\boldsymbol{0}$ 时,$\boldsymbol{d}$ 是 f 在点 $\boldsymbol{x}$ 处的可行下降方向;

(2) $\boldsymbol{d}=\boldsymbol{0}$ 的充要条件是 $\boldsymbol{x}$ 为 K-T 点.

证明 当 $\boldsymbol{d}\neq\boldsymbol{0}$ 时,必有 $\boldsymbol{d}_N\neq\boldsymbol{0}$,如前所述,$\boldsymbol{d}$ 为可行下降方向. 往证(2).

$\boldsymbol{x}$ 为 K-T 点的充要条件是存在乘子 $\boldsymbol{u}=\begin{bmatrix}\boldsymbol{u}_B\\ \boldsymbol{u}_N\end{bmatrix}\geqslant\boldsymbol{0}$ 和 $\boldsymbol{v}$,使得

$$\begin{cases}\begin{bmatrix}\nabla_{x_B}f(\boldsymbol{x})\\ \nabla_{x_N}f(\boldsymbol{x})\end{bmatrix}-\begin{bmatrix}\boldsymbol{B}^{\mathrm{T}}\\ \boldsymbol{N}^{\mathrm{T}}\end{bmatrix}\boldsymbol{v}-\begin{bmatrix}\boldsymbol{u}_B\\ \boldsymbol{u}_N\end{bmatrix}=\begin{bmatrix}\boldsymbol{0}\\ \boldsymbol{0}\end{bmatrix},\\ \boldsymbol{u}_B^{\mathrm{T}}\boldsymbol{x}_B=0,\\ \boldsymbol{u}_N^{\mathrm{T}}\boldsymbol{x}_N=0.\end{cases}\tag{9.3.10}$$

因 $\boldsymbol{x}_B>\boldsymbol{0}$ 且 $\boldsymbol{u}_B\geqslant\boldsymbol{0}$,故 $\boldsymbol{u}_B^{\mathrm{T}}\boldsymbol{x}_B=0$ 当且仅当 $\boldsymbol{u}_B=\boldsymbol{0}$,代入(9.3.10)中第1式得

$$\boldsymbol{v}=(\boldsymbol{B}^{\mathrm{T}})^{-1}\nabla_{x_B}f(\boldsymbol{x}),$$

$$\boldsymbol{u}_N=\nabla_{x_N}f(\boldsymbol{x})-(\boldsymbol{B}^{-1}\boldsymbol{N})^{\mathrm{T}}\nabla_{x_B}f(\boldsymbol{x})=\boldsymbol{r}(\boldsymbol{x}_N).$$

因此,K-T 条件等价于

$$\boldsymbol{r}(\boldsymbol{x}_N)\geqslant\boldsymbol{0},\boldsymbol{r}(\boldsymbol{x}_N)^{\mathrm{T}}\boldsymbol{x}_N=0.\tag{9.3.11}$$

由 $\boldsymbol{d}$ 的定义,$\boldsymbol{d}=\boldsymbol{0}$ 当且仅当 $\boldsymbol{d}_N=\boldsymbol{0}$. 而 $\boldsymbol{d}_N=\boldsymbol{0}$ 当且仅当(9.3.11)式成立. 因此,$\boldsymbol{d}=\boldsymbol{0}$ 的充要条件是 $\boldsymbol{x}$ 为 K-T 点. □

9.3.2 确定一维搜索步长

当迭代点 $\boldsymbol{x}_k\geqslant\boldsymbol{0}$ 时,为保持 $\boldsymbol{x}_{k+1}\geqslant\boldsymbol{0}$,即确定 $\lambda\geqslant0$ 的范围,使得

$$x_j^{(k)}+\lambda d_j^{(k)}\geqslant0,j=1,2,\cdots,n.\tag{9.3.12}$$

当 $d_j^{(k)}>0$ 时,对任意 $\lambda\geqslant0$,(9.3.12)式恒成立. 当 $d_j^{(k)}<0$ 时,应取 $\lambda\leqslant-\dfrac{x_j^{(k)}}{d_j^{(k)}}$,故令

$$\lambda_{\max}=\begin{cases}\min\left\{-\dfrac{x_j^{(k)}}{d_j^{(k)}}\;\middle|\;d_j^{(k)}<0\right\}, & \boldsymbol{d}_k\ngeqslant\boldsymbol{0},\\ \infty, & \boldsymbol{d}_k\geqslant\boldsymbol{0}.\end{cases}\tag{9.3.13}$$

这样,步长的确定归结为一维搜索

$$\min_{0\leqslant\lambda\leqslant\lambda_{\max}}f(\boldsymbol{x}_k+\lambda\boldsymbol{d}_k).$$

算法 9-6(Wolfe 既约梯度法)

Step 1　选取初始数据. 选取初始可行点 $\boldsymbol{x}_0$,给定允许误差 $\varepsilon>0$,令 $k=0$.

Step 2　进行矩阵分解. 从 $\boldsymbol{x}_k$ 中选取 m 个最大分量,对应的变量的指标集记作 J_k,令 $\boldsymbol{B}$ 是由指标属于 J_k 的 $\boldsymbol{A}$ 的列构成的 m 阶方阵,余下的列构成矩阵 $\boldsymbol{N}$.

Step 3　求既约梯度. 计算 $\nabla f(\boldsymbol{x}_k)=\begin{bmatrix}\nabla_{\boldsymbol{x}_B} f(\boldsymbol{x}_k)\\ \nabla_{\boldsymbol{x}_N} f(\boldsymbol{x}_k)\end{bmatrix}$,求

$$\boldsymbol{r}(\boldsymbol{x}_N^{(k)})=\nabla_{\boldsymbol{x}_N} f(\boldsymbol{x}_k)-(\boldsymbol{B}^{-1}\boldsymbol{N})^{\mathrm{T}}\nabla_{\boldsymbol{x}_B} f(\boldsymbol{x}_k).$$

Step 4　构造可行下降方向. 令

$$d_{N_j}^{(k)}=\begin{cases}-x_{N_j}^{(k)} r_j(\boldsymbol{x}_N^{(k)}), & \text{当 } r_j(\boldsymbol{x}_N^{(k)})>0 \text{ 时},\\ -r_j(\boldsymbol{x}_N^{(k)}), & \text{当 } r_j(\boldsymbol{x}_N^{(k)})\leqslant 0 \text{ 时},\end{cases}$$

$$\boldsymbol{d}_B^{(k)}=-\boldsymbol{B}^{-1}\boldsymbol{N}\boldsymbol{d}_N^{(k)}.$$

Step 5　检查终止条件. 若 $\|\boldsymbol{d}_k\|\leqslant\varepsilon$,则停止计算,输出 $\boldsymbol{x}_k$;否则,转 Step 6.

Step 6　进行一维搜索. 求解

$$\min_{0\leqslant\lambda\leqslant\lambda_{\max}} f(\boldsymbol{x}_k+\lambda\boldsymbol{d}_k),$$

得最优解 λ_k,其中 $\lambda_{\max}$按(9.3.13)式计算.

令 $\boldsymbol{x}_{k+1}=\boldsymbol{x}_k+\lambda\boldsymbol{d}_k$, $k:=k+1$,转 Step 2.

例 9.3.1　用 Wolfe 既约梯度法解下列问题

$$\begin{cases}\min & f(\boldsymbol{x})=2x_1^2+2x_2^2-2x_1x_2-4x_1-6x_2;\\ \text{s.t.} & x_1+x_2+x_3=2,\\ & x_1+5x_2+x_4=5,\\ & x_j\geqslant 0, j=1,2,3,4.\end{cases}$$

要求取初始点 $\boldsymbol{x}_0=(0,0,2,5)^{\mathrm{T}}$.

解　$\boldsymbol{A}=\begin{bmatrix}1&1&1&0\\1&5&0&1\end{bmatrix}$, $\nabla f(\boldsymbol{x})=\begin{bmatrix}4x_1-2x_2-4\\4x_1-2x_2-6\\0\\0\end{bmatrix}$.

第一次迭代:

$$J_0=\{3,4\}, \boldsymbol{B}=\begin{bmatrix}1&0\\0&1\end{bmatrix}, \boldsymbol{N}=\begin{bmatrix}1&1\\1&5\end{bmatrix},$$

$$\nabla f(\boldsymbol{x}_0)=(-4,-6,0,0)^{\mathrm{T}}.$$

$$\begin{aligned}\boldsymbol{r}(x_N^{(0)})&=\nabla_{\boldsymbol{x}_N} f(\boldsymbol{x}_0)-(\boldsymbol{B}^{-1}\boldsymbol{N})^{\mathrm{T}}\nabla_{\boldsymbol{x}_B} f(\boldsymbol{x}_0)\\&=\begin{bmatrix}-4\\-6\end{bmatrix}-\left[\begin{bmatrix}1&0\\0&1\end{bmatrix}^{-1}\begin{bmatrix}1&1\\1&5\end{bmatrix}\right]^{\mathrm{T}}\begin{bmatrix}0\\0\end{bmatrix}=\begin{bmatrix}-4\\-6\end{bmatrix},\end{aligned}$$

$$d_N^{(0)}=(4,6)^{\mathrm T},d_B^{(0)}=-B^{-1}Nd_N^{(0)}=(-10,-34)^{\mathrm T},$$

故得 $d_0=(4,6,-10,-34)^{\mathrm T}$ 为可行下降方向. 因 $d_0\neq 0$,沿 d_0 进行一维搜索,求出

$$\lambda_{\max}=\min\left\{\frac{2}{10},\frac{5}{34}\right\}=\frac{5}{34},$$

并求解

$$\min_{0\leqslant\lambda\leqslant\frac{5}{34}} f(x_0+\lambda d_0)=56\lambda^2-52\lambda,$$

得最优解 $\lambda_0=\frac{5}{34}$. 于是,下一迭代点

$$x_1=x_0+\lambda_0 d_0=\left(\frac{10}{17},\frac{15}{17},\frac{9}{17},0\right)^{\mathrm T}.$$

第二次迭代:

$$J_1=\{1,2\},B=\begin{bmatrix}1&1\\1&5\end{bmatrix},N=\begin{bmatrix}1&0\\0&1\end{bmatrix},$$

$$\nabla f(x_1)=\left(-\frac{58}{17},-\frac{62}{17},0,0\right)^{\mathrm T}.$$

$$r(x_N^{(1)})=\begin{bmatrix}0\\0\end{bmatrix}-\left[\begin{bmatrix}1&1\\1&5\end{bmatrix}^{-1}\begin{bmatrix}1&0\\0&1\end{bmatrix}\right]^{\mathrm T}\begin{bmatrix}-\frac{58}{17}\\-\frac{62}{17}\end{bmatrix}=\begin{bmatrix}\frac{57}{17}\\\frac{17}{17}\end{bmatrix},$$

$$d_N^{(1)}=\left(-\frac{513}{289},0\right)^{\mathrm T},d_B^{(1)}=-B^{-1}Nd_N^{(1)}=\left(\frac{2\,565}{1\,156},-\frac{513}{1\,156}\right)^{\mathrm T},$$

故得 $d_1=\left(\frac{2\,565}{1\,156},-\frac{513}{1\,156},-\frac{513}{289},0\right)^{\mathrm T}$ 为可行下降方向. 因 $d_1\neq 0$,沿 d_1 进行一维搜索,求出

$$\lambda_{\max}=\min\left\{\frac{15/17}{513/1\,156},\frac{9/17}{513/289}\right\}=\frac{17}{57},$$

并求解

$$\min_{0\leqslant\lambda\leqslant\frac{17}{57}} f(x_1+\lambda d_1)=\frac{8\,158\,239}{668\,168}\lambda^2-\frac{29\,241}{4\,913}\lambda-\frac{1\,860}{289},$$

得最优解 $\lambda_1=\frac{68}{279}$. 于是,下一迭代点

$$x_2=x_1+\lambda_1 d_1=\left(\frac{35}{31},\frac{24}{31},\frac{3}{31},0\right)^{\mathrm T}.$$

第三次迭代:

$$J_1=\{1,2\},B=\begin{bmatrix}1&1\\1&5\end{bmatrix},N=\begin{bmatrix}1&0\\0&1\end{bmatrix},$$

$$\nabla f(x_2) = \left(-\frac{32}{31}, -\frac{160}{31}, 0, 0 \right).$$

$$r(x_N^{(2)}) = \begin{bmatrix} 0 \\ 0 \end{bmatrix} - \left[\begin{bmatrix} 1 & 1 \\ 1 & 5 \end{bmatrix}^{-1} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \right]^{\mathrm{T}} \begin{bmatrix} -\frac{32}{31} \\ -\frac{160}{31} \end{bmatrix} = \begin{bmatrix} 0 \\ \frac{32}{31} \end{bmatrix},$$

$$d_N^{(2)} = (0,0)^{\mathrm{T}}, d_B^{(2)} = (0,0)^{\mathrm{T}},$$

于是,$d_2 = (0,0,0,0,0)^{\mathrm{T}}$. 因为 $d_2 = \mathbf{0}$,从而停止迭代,$x_2 = \left(\frac{35}{31}, -\frac{24}{31}, \frac{3}{31}, 0 \right)^{\mathrm{T}}$ 为 K-T点,也是最优解. □

图 9.3.1 给出了其迭代过程的图示.

最后指出, Abadie 和 Carpentier (1969)成功地把 Wolfe 既约梯度法推广于求解带非线性等式约束的情形,提出了著名的 GRG 法(generalized reduced gradient method,广义既约梯度法). 数值实例表明,GRG 法是目前求解约束非线性最优化问题的最有效的方法之一(参见文献[3,50]).

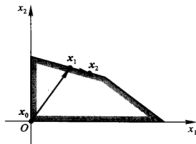


图 9.3.1 Wolfe 既约梯度法迭代进程

9.4 Frank-Wolfe 方法

1956 年,Frank 和 Wolfe 提出了一种求解线性约束问题的算法,其基本思想是将目标函数 $f(x)$ 作线性近似,通过求解线性规划求得可行下降方向,并沿该方向在可行域内作一维搜索. 这种方法又称为近似线性化方法(approximation linearization method).

考虑问题

$$\begin{cases} \min & f(x); \\ \text{s.t.} & Ax = b, \\ & x \geq 0, \end{cases} \tag{9.4.1}$$

其中 A 为 $m \times n$ 行满秩矩阵,$b \in \mathbb{R}^m$,$f: \mathbb{R}^n \to \mathbb{R}$ 为可微函数,并记可行域为

$$S = \{x \mid Ax = b, x \geq 0\}.$$

9.4.1 近似线性化和可行下降方向

设迭代点 $x_k \in S$,利用 Taylor 公式得到 f 在点 x_k 处的线性近似

$$f(x) \approx f(x_k) + \nabla f(x_k)^{\mathrm{T}}(x - x_k). \tag{9.4.2}$$

用上式右边的线性函数近似代替问题(9.4.1)中的目标函数 $f(\boldsymbol{x})$,就得到了与问题(9.4.1)近似的一个线性规划问题

$$\begin{cases} \min & f(\boldsymbol{x}_k)+\nabla f(\boldsymbol{x}_k)^{\mathrm{T}}(\boldsymbol{x}-\boldsymbol{x}_k); \\ \mathrm{s.t.} & \boldsymbol{x}\in S. \end{cases} \tag{9.4.3}$$

去掉目标函数中的常数项,问题(9.4.3)等价于

$$\begin{cases} \min & \nabla f(\boldsymbol{x}_k)^{\mathrm{T}}\boldsymbol{x}; \\ \mathrm{s.t.} & \boldsymbol{x}\in S. \end{cases} \tag{9.4.4}$$

假设此问题存在有限最优解 $\boldsymbol{y}_k$,则由线性规划的基本定理可知,这个最优解可在某个极点上达到.

现在我们讨论问题(9.4.1)和问题(9.4.4)的关系.

定理 9.4.1 设 $f:\mathbb{R}^n\to\mathbb{R}$ 可微,$\boldsymbol{x}_k\in S$,如果 $\boldsymbol{y}_k$ 是问题(9.4.4)的最优解,则

(1) 当 $\nabla f(\boldsymbol{x}_k)^{\mathrm{T}}(\boldsymbol{y}_k-\boldsymbol{x}_k)=0$ 时,$\boldsymbol{x}_k$ 是问题(9.4.1)的 K-T 点;

(2) 当 $\nabla f(\boldsymbol{x}_k)^{\mathrm{T}}(\boldsymbol{y}_k-\boldsymbol{x}_k)\neq 0$ 时,向量

$$\boldsymbol{d}_k=\boldsymbol{y}_k-\boldsymbol{x}_k$$

是 f 在点 $\boldsymbol{x}_k$ 处的可行下降方向.

证明 (1) 设 $\nabla f(\boldsymbol{x}_k)^{\mathrm{T}}(\boldsymbol{y}_k-\boldsymbol{x}_k)=0$,因 $\boldsymbol{y}_k$ 是问题(9.4.4)的最优解,则 $\boldsymbol{x}_k$ 也是问题(9.4.4)的最优解,因此,$\boldsymbol{x}_k$ 是问题(9.4.4)的 K-T 点. 所以,存在乘子向量 $\boldsymbol{u}\geqslant\boldsymbol{0}(\boldsymbol{u}\in\mathbb{R}^n)$ 和 $\boldsymbol{v}\in\mathbb{R}^m$,使

$$\begin{cases} \nabla f(\boldsymbol{x}_k)-\boldsymbol{u}-\boldsymbol{A}^{\mathrm{T}}\boldsymbol{v}=\boldsymbol{0}, \\ \boldsymbol{u}^{\mathrm{T}}\boldsymbol{x}_k=0, \\ \boldsymbol{A}\boldsymbol{x}_k=\boldsymbol{b}, \\ \boldsymbol{x}_k\geqslant\boldsymbol{0}. \end{cases}$$

以上诸式恰为问题(9.4.1)的 K-T 条件,所以 $\boldsymbol{x}_k$ 也是问题(9.4.1)的 K-T 点.

(2) 设 $\nabla f(\boldsymbol{x}_k)^{\mathrm{T}}(\boldsymbol{y}_k-\boldsymbol{x}_k)\neq 0$,则 $\nabla f(\boldsymbol{x}_k)^{\mathrm{T}}(\boldsymbol{y}_k-\boldsymbol{x}_k)<0$,因此,$\boldsymbol{d}_k=\boldsymbol{y}_k-\boldsymbol{x}_k$ 为点 $\boldsymbol{x}_k$ 处的下降方向.

又因 S 为凸集,$\boldsymbol{x}_k\in S$,$\boldsymbol{y}_k\in S$,则连接 $\boldsymbol{x}_k$ 和 $\boldsymbol{y}_k$ 的线段必含于 S,即对每一个 $\lambda\in[0,1]$,有

$$\lambda\boldsymbol{y}_k+(1-\lambda)\boldsymbol{x}_k=\boldsymbol{x}_k+\lambda\boldsymbol{d}_k\in S.$$

故 $\boldsymbol{d}_k$ 是可行方向. 从而 $\boldsymbol{d}_k$ 为可行下降方向. □

9.4.2 Frank-Wolfe 算法及其收敛性

算法 9-7(Frank-Wolfe 近似线性化方法)

Step 1 选取初始数据. 取初始点 $\boldsymbol{x}_0\in S$,给定允许误差 $\varepsilon>0$,令 $k=0$.

Step 2 求解近似线性规划. 求解线性规划问题

$$\begin{cases} \min & \nabla f(\boldsymbol{x}_k)^{\mathrm{T}}\boldsymbol{x}; \\ \text{s.t.} & \boldsymbol{x}\in S, \end{cases}$$

得最优解 $\boldsymbol{y}_k$.

Step 3　构造可行下降方向. 令 $\boldsymbol{d}_k=\boldsymbol{y}_k-\boldsymbol{x}_k$,若 $\|\nabla f(\boldsymbol{x}_k)^{\mathrm{T}}\boldsymbol{d}_k\|\leqslant\varepsilon$,停止计算,输出 $\boldsymbol{x}_k$;否则,$\boldsymbol{d}_k$ 为可行下降方向,转 Step 4.

Step 4　进行一维搜索,求解

$$\min_{0\leqslant\lambda\leqslant1} f(\boldsymbol{x}_k+\lambda\boldsymbol{d}_k),$$

得最优解 λ_k. 令 $\boldsymbol{x}_{k+1}=\boldsymbol{x}_k+\lambda_k\boldsymbol{d}_k$, $k:=k+1$,转 Step 2.

下面给出 Frank－Wolfe 法的收敛定理.

定理 9.4.2　设 $f:\mathbb{R}^n\to\mathbb{R}$ 具有一阶连续偏导数,S 有界,$\boldsymbol{x}_0\in S$, $\{\boldsymbol{x}_k\}$ 是由 Frank－Wolfe 算法产生的迭代点列,则

(1) 当 $\{\boldsymbol{x}_k\}$ 是有穷点列时,其最后一个点是问题(9.4.1)的 K－T 点;

(2) 当 $\{\boldsymbol{x}_k\}$ 是无穷点列时,它必有极限点,并且其任一极限点 $\bar{\boldsymbol{x}}\in S$ 是问题(9.4.1)的 K－T 点;

证明　(1) 若算法在有限步终止,根据终止规则 $\nabla f(\boldsymbol{x}_k)^{\mathrm{T}}(\boldsymbol{y}_k-\boldsymbol{x}_k)=0$ 及定理 9.4.1 知 $\boldsymbol{x}_k$ 为 K－T 点.

(2) 若算法产生一个无穷点列 $\{\boldsymbol{x}_k\}$,则对任意 k,都有

$$\boldsymbol{d}_k=\boldsymbol{y}_k-\boldsymbol{x}_k\neq\boldsymbol{0}.$$

则 $\boldsymbol{d}_k$ 为可行下降方向,即

$$\nabla f(\boldsymbol{x}_k)^{\mathrm{T}}(\boldsymbol{y}_k-\boldsymbol{x}_k)<0. \tag{9.4.5}$$

又从 S 有界得知 $\{\boldsymbol{x}_k\}\subset S$ 是有界点列,因此它必有极限点.

设 $\bar{\boldsymbol{x}}$ 是 $\{\boldsymbol{x}_k\}$ 的极限点,则存在子列 $\{\boldsymbol{x}_k\}_{K_1}\subset\{\boldsymbol{x}_k\}$ $(K_1\subset K=\{0,1,\cdots\})$ 收敛于 $\bar{\boldsymbol{x}}$,即对于 $k\in K_1$,有

$$\lim_{k\to\infty}\boldsymbol{x}_k=\bar{\boldsymbol{x}}.$$

而 S 为闭集,故 $\bar{\boldsymbol{x}}\in S$.

又因 $\{\boldsymbol{x}_{k+1}\}_{K_1}\subset S$ 有界,故它也有极限点 $\bar{\boldsymbol{x}}$,并且存在子列 $\{\boldsymbol{x}_{k+1}\}_{K_2}\subset\{\boldsymbol{x}_{k+1}\}_{K_1}$ $(K_2\subset K_1)$,使得对于 $k\in K_2$ 有

$$\lim_{k\to\infty}\boldsymbol{x}_{k+1}=\bar{\boldsymbol{x}}.$$

由于 S 有界,各近似线性规划问题 $\min\limits_{x\in S}\nabla f(\boldsymbol{x}_k)^{\mathrm{T}}\boldsymbol{x}$ 的最优解 $\boldsymbol{y}_k$ 存在,并且 $\{\boldsymbol{y}_k\}\subset S$,因此 $\{\boldsymbol{y}_k\}_{K_2}\subset S$ 有界,从而必有极限点 $\bar{\boldsymbol{y}}\in S$,故必存在子列 $\{\boldsymbol{y}_k\}_{K_3}$ $(K_3\subset K_2)$,使得对 $k\in K_3$,有

$$\lim_{k\to\infty}\boldsymbol{y}_k=\bar{\boldsymbol{y}}.$$

又 $\boldsymbol{x}_{k+1}=\boldsymbol{x}_k+\lambda_k(\boldsymbol{y}_k-\boldsymbol{x}_k)$, $k\in K$,其中 λ_k 为一维极小化问题

$$\min_{0\le\lambda\le 1} f(\boldsymbol{x}_k+\lambda_k(\boldsymbol{y}_k-\boldsymbol{x}_k))$$

的最优解,故对任意 $\lambda\in[0,1]$ 及 $k\in K$,都有

$$f(\boldsymbol{x}_{k+1})=f(\boldsymbol{x}_k+\lambda_k(\boldsymbol{y}_k-\boldsymbol{x}_k))\leqslant f(\boldsymbol{x}_k+\lambda(\boldsymbol{y}_k-\boldsymbol{x}_k)). \tag{9.4.6}$$

注意到 f 是连续函数,在上式中考虑 $k\in K_3$,并令 $k\to\infty$,得

$$f(\bar{\boldsymbol{x}})\leqslant f(\bar{\boldsymbol{x}}+\lambda(\bar{\boldsymbol{y}}-\bar{\boldsymbol{x}})),\ \forall\lambda\in[0,1], \tag{9.4.7}$$

从(9.4.6)式,由 Taylor 展式,对充分小的 $\lambda>0$ 及任意 $k\in K$,有

$$\begin{aligned}f(\boldsymbol{x}_{k+1})&\leqslant f(\boldsymbol{x}_k+\lambda(\boldsymbol{y}_k-\boldsymbol{x}_k))\\&=f(\boldsymbol{x}_k)+\lambda\nabla f(\boldsymbol{x}_k)^{\mathrm{T}}(\boldsymbol{y}_k-\boldsymbol{x}_k)+o(\lambda\|\boldsymbol{y}_k-\boldsymbol{x}_k\|).\end{aligned}$$

根据(9.4.5)式,得

$$f(\boldsymbol{x}_{k+1})<f(\boldsymbol{x}_k),k\in K,$$

这说明 $\{f(\boldsymbol{x}_k)\}$ 是严格单调下降的,又因 f 连续,S 有界,故 $\{f(\boldsymbol{x}_k)\}$ 是有界数列,所以它必有极限,即

$$\min_{k\to\infty} f(\boldsymbol{x}_k)=\bar{f}.$$

根据 f 的连续性,并考虑 $k\in K_2$,那么有

$$f(\bar{\boldsymbol{x}})=\min_{k\to\infty} f(\boldsymbol{x}_k)=\bar{f}=\min_{k\to\infty} f(\boldsymbol{x}_{k+1})=f(\bar{\boldsymbol{x}}). \tag{9.4.8}$$

往证

$$\nabla f(\bar{\boldsymbol{x}})^{\mathrm{T}}(\bar{\boldsymbol{y}}-\bar{\boldsymbol{x}})=0. \tag{9.4.9}$$

用反证法.若不然,假设 $\nabla f(\bar{\boldsymbol{x}})^{\mathrm{T}}(\bar{\boldsymbol{y}}-\bar{\boldsymbol{x}})\neq 0$,因注意到 $\nabla f(\boldsymbol{x})$ 连续并考虑 $k\in K_3$,由(9.4.5)式得

$$\nabla f(\bar{\boldsymbol{x}})^{\mathrm{T}}(\bar{\boldsymbol{y}}-\bar{\boldsymbol{x}})\leqslant 0,$$

从而必有

$$\nabla f(\bar{\boldsymbol{x}})^{\mathrm{T}}(\bar{\boldsymbol{y}}-\bar{\boldsymbol{x}})<0,$$

所以,$\bar{\boldsymbol{d}}=\bar{\boldsymbol{y}}-\bar{\boldsymbol{x}}$ 是点 $\bar{\boldsymbol{x}}$ 处的下降方向,即存在 $\delta>0$,使

$$f(\bar{\boldsymbol{x}}+\lambda\bar{\boldsymbol{d}})<f(\bar{\boldsymbol{x}}),\ \forall\lambda\in(0,\delta).$$

再由(9.4.7)式知 $f(\bar{\boldsymbol{x}})<f(\bar{\boldsymbol{x}})$,但这与(9.4.8)式矛盾,所以(9.4.9)式成立.

另一方面,由 $\boldsymbol{y}_k$ 的含义知,对一切 $\boldsymbol{x}\in S$,都有

$$\nabla f(\boldsymbol{x}_k)^{\mathrm{T}}\boldsymbol{y}_k\leqslant\nabla f(\boldsymbol{x}_k)^{\mathrm{T}}\boldsymbol{x}.$$

考虑 $k\in K_3$,并令 $k\to\infty$,有

$$\nabla f(\bar{\boldsymbol{x}})^{\mathrm{T}}\bar{\boldsymbol{y}}\leqslant\nabla f(\bar{\boldsymbol{x}})^{\mathrm{T}}\boldsymbol{x},$$

且 $\bar{\boldsymbol{x}}\in S$,故 $\bar{\boldsymbol{y}}$ 是线性规划问题

$$\begin{cases}\min & \nabla f(\bar{\boldsymbol{x}})^{\mathrm{T}}\boldsymbol{x};\\ \text{s.t.} & \boldsymbol{x}\in S\end{cases}$$

的最优解.

根据这一事实和(9.4.9)式,应用定理 9.4.1 得知,$\bar{\boldsymbol{x}}$ 为问题(9.4.1)的 K -

T 点. □

Frank - Wolfe 算法是一种可行方向法,在每次迭代中,搜索方向总是指向某个极点,并且当迭代点接近最优解时,搜索方向与目标函数的梯度趋于正交,因此算法收敛速度比较慢. 但该方法把求解非线性最优化问题转化为求解一系列线性规划问题,而且各线性规划具有相同的约束条件,因而该方法在实际应用中仍然是一种有用的算法.

例 9.4.1 用 Frank - Wolfe 法求解下列问题:

$$\begin{cases} \min & f(\boldsymbol{x}) = 2x_1^2 + 2x_2^2 - 2x_1x_2 - 4x_1 - 6x_2; \\ \text{s.t.} & x_1 + x_2 \leqslant 2, \\ & x_1 + 5x_2 \leqslant 5, \\ & x_1, x_2 \geqslant 0, \end{cases} \tag{9.4.10}$$

要求取初始可行点 $\boldsymbol{x}_0 = (0,0)^{\mathrm{T}}$.

解 $\nabla f(\boldsymbol{x}) = \begin{bmatrix} 4x_1 - 2x_2 - 4 \\ 4x_1 - 2x_2 - 6 \end{bmatrix}$, $\nabla f(\boldsymbol{x}_0) = (-4, -6)^{\mathrm{T}}$.

第一次迭代:

解近似线性规划问题

$$\begin{cases} \min & \nabla f(\boldsymbol{x}_0)^{\mathrm{T}}\boldsymbol{x} = -4x_1 - 6x_2; \\ \text{s.t.} & x_1 + x_2 \leqslant 2, \\ & x_1 + 5x_2 \leqslant 5, \\ & x_1, x_2 \geqslant 0, \end{cases}$$

得最优解 $\boldsymbol{y}_0 = \left(\frac{5}{4}, \frac{3}{4}\right)^{\mathrm{T}}$. 因 $\nabla f(\boldsymbol{x}_0)^{\mathrm{T}}(\boldsymbol{y}_0 - \boldsymbol{x}_0) = -\frac{19}{2} \neq 0$, 故 $\boldsymbol{d}_0 = \boldsymbol{y}_0 - \boldsymbol{x}_0 = \left(\frac{5}{4}, \frac{3}{4}\right)^{\mathrm{T}}$ 为可行下降方向.

求解

$$\min_{0 \leqslant \lambda \leqslant 1} f(\boldsymbol{x}_0 + \lambda \boldsymbol{d}_0) = \frac{19}{8}\lambda^2 - \frac{19}{2}\lambda,$$

得最优解 $\lambda_0 = 1$. 于是,下一迭代点为

$$\boldsymbol{x}_1 = \boldsymbol{x}_0 + \lambda_0 \boldsymbol{d}_0 = \left(\frac{5}{4}, \frac{3}{4}\right)^{\mathrm{T}}.$$

第二次迭代:

$\nabla f(\boldsymbol{x}_1) = \left(-\frac{1}{2}, -\frac{11}{2}\right)^{\mathrm{T}}$,解近似线性规划问题

$$\begin{cases} \min & \nabla f(\boldsymbol{x}_1)^{\mathrm{T}}\boldsymbol{x} = -\dfrac{1}{2}x_1 - \dfrac{11}{2}x_2; \\ \text{s.t.} & x_1 + x_2 \leqslant 2, \\ & x_1 + 5x_2 \leqslant 5, \\ & x_1, x_2 \geqslant 0, \end{cases}$$

得最优解 $\boldsymbol{y}_1 = (0,1)^{\mathrm{T}}$. 因 $\nabla f(\boldsymbol{x}_1)^{\mathrm{T}}(\boldsymbol{y}_1 - \boldsymbol{x}_1) = -\dfrac{3}{4} \neq 0$, 故 $\boldsymbol{d}_1 = \boldsymbol{y}_1 - \boldsymbol{x}_1 = \left(-\dfrac{5}{4}, \dfrac{1}{4}\right)^{\mathrm{T}}$ 为可行下降方向.

求解

$$\min_{0 \leqslant \lambda \leqslant 1} f(\boldsymbol{x}_1 + \lambda \boldsymbol{d}_1) = \frac{31}{8}\lambda^2 - \frac{3}{4}\lambda - \frac{57}{8},$$

得最优解 $\lambda = \dfrac{3}{31}$. 于是,下一迭代点为

$$\boldsymbol{x}_2 = \boldsymbol{x}_1 + \lambda_1 \boldsymbol{d}_1 = \left(\frac{35}{31}, \frac{24}{31}\right)^{\mathrm{T}}.$$

第三次迭代:

$\nabla f(\boldsymbol{x}_2) = \left(-\dfrac{32}{31}, -\dfrac{160}{31}\right)^{\mathrm{T}}$,解近似线性规划问题

$$\begin{cases} \min & \nabla f(\boldsymbol{x}_2)^{\mathrm{T}}\boldsymbol{x} = -\dfrac{32}{31}x_1 - \dfrac{160}{31}x_2; \\ \text{s.t.} & x_1 + x_2 \leqslant 2, \\ & x_1 + 5x_2 \leqslant 5, \\ & x_1, x_2 \geqslant 0, \end{cases}$$

得最优解(不唯一) $\boldsymbol{y}_1 = \left(\tau, 1 - \dfrac{\tau}{5}\right)^{\mathrm{T}}$ $\left(0 \leqslant \tau \leqslant \dfrac{5}{4}\right)$.

因 $\nabla f(\boldsymbol{x}_2)^{\mathrm{T}}(\boldsymbol{y}_2 - \boldsymbol{x}_2) = 0$,停止迭代,得 $\boldsymbol{x}_2 = \left(\dfrac{35}{31}, \dfrac{24}{31}\right)^{\mathrm{T}}$ 为问题(9.4.10)的 K-T 点,也是最优解. □

图 9.4.1 表示了其迭代过程.

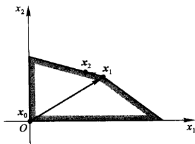


图 9.4.1 Frank-Wolfe 法迭代进程

习 题 九

1. 用 Zoutendijk 方法求解下列问题:

$$\begin{cases}\min & x_1^2+4x_2^2-34x_1-32x_2;\\ \text{s.t.} & 2x_1+x_2\leqslant 6,\\ & x_2\leqslant 2,\\ & x_1,x_2\geqslant 0,\end{cases}$$

要求取初始可行点 $\boldsymbol{x}_0=(1,2)^{\mathrm{T}}$.

2. 考虑以下问题

$$\begin{cases}\min & x_1^2+2x_2^2+x_1x_2-6x_1-2x_2-12x_3;\\ \text{s.t.} & x_1-2x_2\geqslant -3,\\ & x_1,x_2,x_3\geqslant 0,\\ & x_1+x_2+x_3=2,\end{cases}$$

试用梯度投影法求出在点 $\hat{\boldsymbol{x}}=(1,1,0)^{\mathrm{T}}$ 处的一个可行下降方向.

3. 用梯度投影法求解下列问题

$$\begin{cases}\min & x_1^2+x_2^2-6x_1+2x_2;\\ \text{s.t.} & x_1+2x_2\leqslant 2,\\ & x_1,x_2\geqslant 0,\end{cases}$$

要求取初始可行点 $\boldsymbol{x}_0=(0,1)^{\mathrm{T}}$.

4. 用梯度投影法求解下列问题

$$\begin{cases}\min & (1-x_1)^2-10(x_2-x_1)+x_1^2-2x_1x_2+e^{-x_1-x_2};\\ \text{s.t.} & 2x_1+5x_2\leqslant 25,\\ & -x_1+2x_2\leqslant 8,\\ & x_1,x_2\geqslant 0.\end{cases}$$

5. 用 Wolfe 既约梯度法求解下列问题

$$\begin{cases}\min & 2x_1^2+x_2^2;\\ \text{s.t.} & x_1-x_2+x_3=2,\\ & -2x_1+x_2+x_4=1,\\ & x_1,x_2,x_3,x_4\geqslant 0,\end{cases}$$

要求取初始可行点 $\boldsymbol{x}_0=(1,3,4,0)^{\mathrm{T}}$,

6. 用 Frank - Wolfe 法求解下列问题

$$\begin{cases}\min & -x_1x_2;\\ \text{s.t.} & x_1+x_2\leqslant 2,\\ & x_1,x_2\geqslant 0,\end{cases}$$

要求取初始可行点 $\boldsymbol{x}_0=(2,0)^{\mathrm{T}}$.

7. 设 $\bar{\boldsymbol{x}}$ 是下列问题

$$\begin{cases}\min & f(\boldsymbol{x});\\ \text{s.t.} & g_i(\boldsymbol{x})\geqslant 0, i=1,2,\cdots,m,\\ & h_j(\boldsymbol{x})\geqslant 0, j=1,2,\cdots,l\end{cases}$$

的可行点,记 $I(\bar{\boldsymbol{x}})=\{i\mid g_i(\bar{\boldsymbol{x}})=0, i=1,2,\cdots,m\}$,证明 $\bar{\boldsymbol{x}}$ 是 K－T 点的充要条件是下列问题

$$\begin{cases}\min & \nabla f(\boldsymbol{x})^{\mathrm{T}}\boldsymbol{d};\\ \text{s.t.} & \nabla g_i(\bar{\boldsymbol{x}})^{\mathrm{T}}\boldsymbol{d}\geqslant 0, i\in I(\bar{\boldsymbol{x}}),\\ & \nabla h_j(\bar{\boldsymbol{x}})^{\mathrm{T}}\boldsymbol{d}=0, j=1,2,\cdots,l,\\ & -1\leqslant d_j\leqslant 1, j=1,2,\cdots,n\end{cases}$$

的目标函数最优值为 0.

第十章　罚函数法与广义乘子法

求解约束最优化问题，一种途径是在可行域内寻找使目标函数值下降的迭代点列，即第九章介绍的各种可行方向法. 但是，这类方法对于带非线性约束的最优化问题求解效果一般都不甚理想. 因此，本章将介绍求解约束最优化问题的另一种途径，即利用问题的目标函数和约束函数构造新的目标函数——罚函数（penalty function），把约束最优化问题转化为相应的罚函数的无约束最优化问题来求解.

本章先介绍直接利用约束最优化问题的目标函数和约束函数构造罚函数的两种罚函数法（penalty function method）. 然后，研究在改进罚函数法的过程中产生并发展起来的、借助 Lagrange 乘子来构造罚函数的广义乘子法（generlized multiplier method）.

10.1　外罚函数法

本章考虑约束非线性最优化问题

$$\begin{cases} \min & f(\boldsymbol{x}); \\ \text{s.t.} & g_i(\boldsymbol{x}) \geqslant 0, i=1,2,\cdots,m, \\ & h_j(\boldsymbol{x})=0, j=1,2,\cdots,l, \end{cases} \tag{10.1.1}$$

其中 $f(\boldsymbol{x}), g_i(\boldsymbol{x})\ (i=1,2,\cdots,m)$ 和 $h_j(\boldsymbol{x})\ (j=1,2,\cdots,l)$ 都是定义在 $\mathbb{R}^n$ 上的实值连续函数. 记问题(10.1.1)的可行域为 S.

问题(10.1.1)的罚函数是指利用目标函数 $f(\boldsymbol{x})$ 和约束函数 $g_i(\boldsymbol{x})$ 及 $h_j(\boldsymbol{x})$ 所构造的、具有“惩罚性质”的辅助函数

$$F(\boldsymbol{x}) = \widetilde{F}(f(\boldsymbol{x}), g_i(\boldsymbol{x}), h_j(\boldsymbol{x})).$$

所谓“惩罚性质”就是要求 $\boldsymbol{x}\in S$ 当且仅当 $F(\boldsymbol{x})=f(\boldsymbol{x})$；而 $\boldsymbol{x}\notin S$ 时，$F(\boldsymbol{x}) > f(\boldsymbol{x})$，并且 $F(\boldsymbol{x})-f(\boldsymbol{x})$ 随着 $\boldsymbol{x}$ 到 S 的距离的增大而增大.

例如，对于等式约束问题

$$\begin{cases} \min & f(\boldsymbol{x}); \\ \text{s.t.} & h_j(\boldsymbol{x})=0, j=1,2,\cdots,l, \end{cases} \tag{10.1.2}$$

可定义罚函数

$$F_1(\boldsymbol{x},\sigma)=f(\boldsymbol{x})+\sigma\sum_{j=1}^{l}h_j^2(\boldsymbol{x}),$$

其中参数 σ 是一很大的正常数. 这样,就把问题(10.1.2)转化为无约束问题

$$\min_{\boldsymbol{x}\in\mathbb{R}^n}F_1(\boldsymbol{x},\sigma).\tag{10.1.3}$$

显然,问题(10.1.3)的最优解必使得所有 $h_j(\boldsymbol{x})\,(j=1,2,\cdots,l)$ 都接近 0,否则,罚函数 $F_1(\boldsymbol{x},\sigma)$ 的第二项将是很大的正数,与最优解取到极小值矛盾. 因此,通过求解无约束问题(10.1.3)能够得到等式约束问题(10.1.2)的近似解.

对于不等式约束问题

$$\begin{cases}\min & f(\boldsymbol{x});\\ \text{s.t.} & g_i(\boldsymbol{x})\geqslant 0,\ i=1,2,\cdots,m,\end{cases}\tag{10.1.4}$$

辅助函数的形式与等式约束情形不同,我们定义罚函数

$$F_2(\boldsymbol{x},\sigma)=f(\boldsymbol{x})+\sigma\sum_{i=1}^{m}[\max\{0,-g_i(\boldsymbol{x})\}]^2,$$

其中参数 σ 是很大的正数. 这样,可将问题(10.1.4)转化为无约束问题

$$\min_{\boldsymbol{x}\in\mathbb{R}^n}F_2(\boldsymbol{x},\sigma).\tag{10.1.5}$$

同理,问题(10.1.5)的最优解也必使得所有 $g_i(\boldsymbol{x})\,(i=1,2,\cdots,m)$ 大于 0 或接近 0. 因此,通过求解问题(10.1.5)可求得不等式约束问题(10.1.4)的近似解.

我们把上述思想加以推广,对于一般的约束最优化问题(10.1.1),可定义罚函数

$$F(\boldsymbol{x},\sigma)=f(\boldsymbol{x})+\sigma P(\boldsymbol{x}),$$

其中参数 σ 是很大的正数,$P(\boldsymbol{x})$ 具有下列形式:

$$P(\boldsymbol{x})=\sum_{i=1}^{m}\varphi[g_i(\boldsymbol{x})]+\sum_{j=1}^{l}\psi[h_j(\boldsymbol{x})],$$

φ 和 ψ 是满足下列条件的实值连续函数:

$$\begin{cases}\varphi(y)=0, & \text{当 } y\geqslant 0 \text{ 时},\\ \varphi(y)>0, & \text{当 } y<0 \text{ 时};\end{cases}$$

$$\begin{cases}\psi(y)=0, & \text{当 } y=0 \text{ 时},\\ \psi(y)>0, & \text{当 } y\neq 0 \text{ 时}.\end{cases}$$

函数 φ 和 ψ 的典型取法为

$$\varphi(y)=[\max\{0,-y\}]^\alpha,$$

$$\psi(y)=|y|^\beta,$$

其中 $\alpha\geqslant 1$ 和 $\beta\geqslant 1$ 是给定的常数,通常取作 1 或 2.

这样,可将约束问题(10.1.1)转化为无约束问题

$$\min_{\boldsymbol{x}\in\mathbb{R}^n}F(\boldsymbol{x},\sigma)=f(x)+\sigma P(\boldsymbol{x}),\tag{10.1.6}$$

其中 σ 是很大的正数，$P(\boldsymbol{x})$ 是连续函数.

由罚函数 $F(\boldsymbol{x},\sigma)$ 的定义可知，目标函数附加的权项 $\sigma P(\boldsymbol{x})$ 反映了对落在可行域外(即不满足约束条件)的点 $\boldsymbol{x}$ 的一种惩罚，其作用是在极小化过程中迫使迭代点靠近可行域. 显然，问题(10.1.6)的最优解也必使得所有 $g_i(\boldsymbol{x})$ $(i=1,2,\cdots,m)$ 大于 0 或接近 0，而所有 $h_j(\boldsymbol{x})$ $(j=1,2,\cdots,l)$ 都接近 0. 因此，求解问题(10.1.6)能够得到一般约束问题(10.1.1)的近似解，而且 σ 越大，近似程度越好. 通常将 $\sigma P(\boldsymbol{x})$ 称为罚项，参数 σ 称为罚因子(penalty factor).

例 10.1.1 试说明罚函数对求解下面的约束问题所起的作用：

$$\begin{cases} \min & x_1^2+x_2^2; \\ \text{s.t.} & x_2=1. \end{cases} \tag{10.1.7}$$

解 首先定义罚函数

$$F(\boldsymbol{x},\sigma)=x_1^2+x_2^2+\sigma(x_2-1)^2.$$

通过求解无约束问题

$$\min_{\boldsymbol{x}\in\mathbb{R}^n} F(\boldsymbol{x},\sigma)=x_1^2+x_2^2+\sigma(x_2-1)^2, \tag{10.1.8}$$

可求得约束问题(10.1.7)的近似解，我们用解析方法解问题(10.1.8). 根据罚函数 $F(\boldsymbol{x},\sigma)$ 的定义可知，使得 $F(\boldsymbol{x},\sigma)$ 有极小值的必要条件为

$$\begin{cases} \dfrac{\partial F}{\partial x_1}=2x_1=0, \\ \dfrac{\partial F}{\partial x_2}=2x_2+2\sigma(x_2-1)=0. \end{cases}$$

得到问题(10.1.8)的最优解 $\hat{\boldsymbol{x}}_\sigma=(\hat{x}_1^{(\sigma)},\hat{x}_2^{(\sigma)})^{\mathrm{T}}$，其中

$$\hat{x}_1^{(\sigma)}=0,\hat{x}_2^{(\sigma)}=\frac{\sigma}{1+\sigma}.$$

从表 10.1.1 中可以看出，当参数 σ 从 0 增加到 ∞ 时，$\hat{\boldsymbol{x}}_\sigma$ 从无约束解向约束解变化. □

表 10.1.1 σ 值对应的解

σ	0	1	2	5	10	∞
$\hat{x}_2^{(\sigma)}$	0	0.5	0.67	0.83	0.91	1

从图 10.1.1 中我们还发现，随着 σ 值的不断增大，无约束问题(10.1.8)的最优解 $\hat{\boldsymbol{x}}_\sigma$ 在约束问题(10.1.7)的可行域 $\{(x_1,x_2)^{\mathrm{T}} \mid x_2=1\}$ 的外部以高度非线性方式向可行域边界靠拢；并且 σ 越大，$\hat{\boldsymbol{x}}_\sigma$ 越靠近问题(10.1.7)的最优解 $\bar{\boldsymbol{x}}=(0,1)^{\mathrm{T}}$，当 $\sigma\to\infty$ 时，$\hat{\boldsymbol{x}}_\sigma\to\bar{\boldsymbol{x}}$. 那么，是否任何一个约束问题与其对应的无约束问题之间都存在这种联系呢？下面的定理给予了肯定回答.

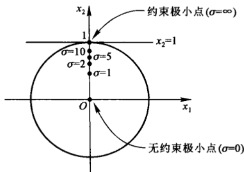


图 10.1.1　例 10.1.1 解的迭代进程

定理 10.1.1　设对于某个给定的 $\sigma>0$,无约束问题(10.1.6)的最优解为 $\hat{x}_\sigma$,S 为约束问题(10.1.1)的可行域,那么

(1) 若 $\hat{x}_\sigma\in S$,则 $\hat{x}_\sigma$ 必为约束问题(10.1.1)的最优解;

(2) 对于任何 $\sigma_1>\sigma$,无约束问题

$$\min_{x\in\mathbb{R}^n} F(x,\sigma_1). \tag{10.1.9}$$

存在最优解 $\hat{x}_{\sigma_1}$,并且若 $\hat{x}_\sigma\in S$,则 $\hat{x}_{\sigma_1}\in S$,即 $\hat{x}_{\sigma_1}$也为问题(10.1.1)的最优解.

证明　(1) 由罚函数 $F(x,\sigma)$的定义,有

$$F(x,\sigma)=f(x),\ \forall x\in S,$$

又注意到 $\hat{x}_\sigma\in S$,所以对一切 $x\in S$,总有

$$f(\hat{x}_\sigma)=F(\hat{x}_\sigma,\sigma)=\min_{x\in\mathbb{R}^n}F(x,\sigma)\leqslant F(x,\sigma)=f(x),$$

即 $\hat{x}_\sigma$ 为问题(10.1.1)的最优解.

(2) 先证:对任何 $\sigma_1>\sigma$,问题(10.1.9)存在最优解.因为 $\hat{x}_\sigma$ 为问题(10.1.6)的最优解且 $\sigma_1>\sigma$,所以

$$F(\hat{x}_\sigma,\sigma)\leqslant F(x,\sigma)\leqslant F(x,\sigma_1),\ \forall x\in\mathbb{R}^n,$$

即 $F(x,\sigma_1)$在$\mathbb{R}^n$ 上有下界 $F(\hat{x}_\sigma,\sigma)$,这表明问题(10.1.9)存在最优解.

再用反证法证明:当 $\hat{x}_\sigma\in S$ 时,问题(10.1.9)的最优解 $\hat{x}_{\sigma_1}\in S$.若不然,$\hat{x}_{\sigma_1}\notin S$,则 $P(\hat{x}_{\sigma_1})>0$,从而由问题(10.1.6)的最优解 $\hat{x}_\sigma\in S$ 和 $\sigma_1>\sigma$ 知,

$$\begin{aligned}f(\hat{x}_\sigma)=F(\hat{x}_\sigma,\sigma)&\leqslant F(\hat{x}_{\sigma_1},\sigma)=f(\hat{x}_{\sigma_1})+\sigma P(\hat{x}_{\sigma_1})\\&<f(\hat{x}_{\sigma_1})+\sigma_1P(\hat{x}_{\sigma_1})=F(\hat{x}_{\sigma_1},\sigma_1).\end{aligned} \tag{10.1.10}$$

另一方面,由 $\hat{x}_{\sigma_1}$为问题(10.1.9)的最优解及 $\hat{x}_\sigma\in S$ 知,

$$F(\hat{x}_{\sigma_1},\sigma_1)\leqslant F(\hat{x}_\sigma,\sigma_1)=f(\hat{x}_\sigma)+\sigma_1P(\hat{x}_\sigma)=f(\hat{x}_\sigma),$$

此与(10.1.10)式矛盾.因此 $\hat{x}_{\sigma_1}\in S$,从而由(1) 知 $\hat{x}_{\sigma_1}$为问题(10.1.1)的最优解. □

这个定理说明,只要求出无约束问题(10.1.6)的最优解 $\hat{x}_\sigma$,且 $\hat{x}_\sigma\in S$,就相

当于求出了原约束问题(10.1.1)的最优解. 若 $\hat{x}_\sigma \notin S$,此时应增大罚因子 σ 的值,再重复上面的讨论.

在实际计算中,为了利用数值解法,问题(10.1.6)中 σ 的值必须取定且适中. 若 σ 太小,则无约束问题的最优解可能远离约束问题的最优解,计算效率低;若 σ 过大,则人为地给无约束问题的求解增加了计算上的困难. 因此,一般的策略是取定数列 $\{\sigma_k\}$ 满足

$$0<\sigma_1<\sigma_2<\cdots<\sigma_k<\cdots,\lim_{k\to\infty}\sigma_k=\infty,$$

并求解无约束问题

$$\min_{x\in\mathbb{R}^n}F(\boldsymbol{x},\sigma_k).$$

若对应于 σ_k 的罚函数的最优解 $\hat{x}_{\sigma_k}\notin S$,并简记 $\hat{x}_{\sigma_k}$ 为 $\boldsymbol{x}_k$,则可以证明:$\boldsymbol{x}_k$ 与可行域 S 的距离将随着 k 的增大而减小,并且序列 $\{\boldsymbol{x}_k\}$ 收敛于约束问题的最优解(见定理 10.1.3). 这种通过求解一系列无约束问题来获得约束问题最优解的方法统称为序列无约束极小化方法(sequential unconstrained minimization technique),简称 SUMT 法.

显然,上述罚函数的最优解序列 $\{\boldsymbol{x}_k\}$ 是从可行域 S 的外部逼近约束问题(10.1.1)的最优解的. 因此,上述利用罚函数生成一系列外点逼近该约束问题最优解的方法称为外罚函数法(exterior penalty function method)或 SUMT 外点法.

算法 10-1(外罚函数法)

Step 1 选取初始数据. 给定初始点 $\boldsymbol{x}_0$,初始罚因子 $\sigma_1>0$,放大系数 $\alpha>1$,允许误差 $\varepsilon>0$,令 $k=1$.

Step 2 求解无约束问题. 以 $\boldsymbol{x}_{k-1}$ 为初始点,求解无约束问题

$$\min_{x\in\mathbb{R}^n}F(\boldsymbol{x},\sigma_k)=f(\boldsymbol{x})+\sigma_kP(\boldsymbol{x}),\tag{10.1.11}$$

设其最优解为 $\boldsymbol{x}_k$.

Step 3 检查是否满足终止准则. 若 $\sigma_kP(\boldsymbol{x}_k)<\varepsilon$,则迭代终止,$\boldsymbol{x}_k$ 为约束问题(10.1.1)的近似最优解;否则,令 $\sigma_{k+1}=\alpha\sigma_k$,$k:=k+1$,返回 Step 2.

例 10.1.2 用外罚函数法求解约束最优化问题

$$\begin{cases}\min & f(\boldsymbol{x})=\dfrac{1}{2}x_1^2+\dfrac{1}{6}x_2^2;\\ \text{s.t.} & x_1+x_2=1,\end{cases}\tag{10.1.12}$$

其中 $\boldsymbol{x}=(x_1,x_2)^{\mathrm{T}}$,要求选取 $\sigma_1=0.05$,$\alpha=2$,$\varepsilon=10^{-5}$.

解 引入罚函数

$$F(\boldsymbol{x},\sigma_k)=\left(\frac{1}{2}x_1^2+\frac{1}{6}x_2^2\right)+\sigma_k(x_1+x_2-1)^2,$$

则问题(10.1.12)相应的一系列无约束最优化问题为

$$\min_{x \in \mathbb{R}^n} F(\boldsymbol{x}, \sigma_k).$$

其中 $\sigma_1 = 0.05, \sigma_{k+1} = 2\sigma_k, k = 1, 2, \cdots$.

用解析方法求解上述无约束问题,令

$$\nabla F(\boldsymbol{x}, \sigma_k) = \boldsymbol{0},$$

可得最优解

$$\boldsymbol{x}_k = (x_1^{(k)}, x_2^{(k)})^{\mathrm{T}} = \left(\frac{2\sigma_k}{1+8\sigma_k}, \frac{6\sigma_k}{1+8\sigma_k}\right)^{\mathrm{T}}.$$

同时,有

$$\sigma_k P(\boldsymbol{x}_k) = \frac{\sigma_k}{(1+8\sigma_k)^2}.$$

依次对 $k = 1, 2, \cdots$,用上述公式计算 $\boldsymbol{x}_k$ 和 $\sigma_k P(\boldsymbol{x}_k)$,结果如表 10.1.2 所示.

表 10.1.2 $\boldsymbol{x}_k$ 与 $\sigma_k P(\boldsymbol{x}_k)$ 的关系

k	$\boldsymbol{x}_k$	$\sigma_k P(\boldsymbol{x}_k)$	k	$\boldsymbol{x}_k$	$\sigma_k P(\boldsymbol{x}_k)$
1	$(0.0714, 0.2142)^{\mathrm{T}}$	2.551×10^{-2}	9	$(0.2475, 0.7427)^{\mathrm{T}}$	1.197×10^{-3}
2	$(0.1111, 0.3333)^{\mathrm{T}}$	3.086×10^{-2}	10	$(0.2487, 0.7463)^{\mathrm{T}}$	6.042×10^{-4}
3	$(0.1538, 0.4615)^{\mathrm{T}}$	2.958×10^{-2}	11	$(0.2493, 0.7481)^{\mathrm{T}}$	3.036×10^{-4}
4	$(0.1904, 0.5714)^{\mathrm{T}}$	2.267×10^{-2}	12	$(0.2496, 0.7490)^{\mathrm{T}}$	1.522×10^{-4}
5	$(0.2162, 0.6486)^{\mathrm{T}}$	1.461×10^{-2}	13	$(0.2498, 0.7494)^{\mathrm{T}}$	7.619×10^{-5}
6	$(0.2313, 0.6956)^{\mathrm{T}}$	8.400×10^{-3}	14	$(0.2499, 0.7497)^{\mathrm{T}}$	3.812×10^{-5}
7	$(0.2406, 0.7218)^{\mathrm{T}}$	4.522×10^{-3}	15	$(0.2499, 0.7498)^{\mathrm{T}}$	1.907×10^{-5}
8	$(0.2452, 0.7356)^{\mathrm{T}}$	4.305×10^{-3}	16	$(0.2499, 0.7499)^{\mathrm{T}}$	9.536×10^{-6}

由迭代终止条件

$$\sigma_k P(\boldsymbol{x}_k) = \frac{\sigma_k}{(1+8\sigma_k)^2} < 10^{-5},$$

可得约束问题(10.1.12)的近似最优解(保留 4 位有效数字)

$$\boldsymbol{x}_{16} = (0.2499, 0.7499)^{\mathrm{T}}.$$

并可注意到每一迭代点 $\boldsymbol{x}_k = (x_1^{(k)}, x_2^{(k)})^{\mathrm{T}}$ 满足

$$x_1^{(k)} + x_2^{(k)} = \frac{8\sigma_k}{1+8\sigma_k} < 1,$$

这说明,迭代点列 $\{\boldsymbol{x}_k\}$ 是从约束问题(10.1.12)的可行域的外部趋向位于边界

上的最优解$\bar{x}=(0.25,0.75)^T$. □

为了讨论外罚函数法的收敛性,我们首先证明下面的引理.

引理 10.1.2 设$0<\sigma_k<\sigma_{k+1}$,x_k和x_{k+1}分别是取罚因子σ_k及σ_{k+1}时的无约束问题(10.1.11)的最优解,则下列各式成立:

(1) $F(x_k,\sigma_k)\leqslant F(x_{k+1},\sigma_{k+1})$;

(2) $P(x_k)\geqslant P(x_{k+1})$;

(3) $f(x_k)\leqslant f(x_{k+1})$.

证明 (1) 因为x_k是问题(10.1.11)取罚因子σ_k时的最优解,所以

$$\begin{aligned}F(x_k,\sigma_k)&\leqslant F(x_{k+1},\sigma_k)=f(x_{k+1})+\sigma_kP(x_{k+1})\\&\leqslant f(x_{k+1})+\sigma_{k+1}P(x_{k+1})=F(x_{k+1},\sigma_{k+1}).\end{aligned}$$

(2) 由于x_k和x_{k+1}分别是问题(10.1.6)取罚因子σ_k和σ_{k+1}时的最优解,因此有

$$f(x_k)+\sigma_kP(x_k)\leqslant f(x_{k+1})+\sigma_kP(x_{k+1}),\qquad(10.1.13)$$

$$f(x_{k+1})+\sigma_{k+1}P(x_{k+1})\leqslant f(x_k)+\sigma_{k+1}P(x_k).$$

将上面两式的两端分别相加,经整理后可得

$$(\sigma_{k+1}-\sigma_k)[P(x_k)-P(x_{k+1})]\geqslant0,$$

又$\sigma_{k+1}-\sigma_k>0$,所以得到$P(x_k)\geqslant P(x_{k+1})$.

(3) 利用(10.1.13)式和结论(2),显然结论成立. □

由上述引理可知,如果迭代不终止,那么$\{f(x_k)\}$和$\{F(x_k,\sigma_k)\}$为单调递增数列,$\{P(x_k)\}$为单调递减数列.

下面给出外罚函数法的收敛性定理.

定理 10.1.3 设约束最优化问题(10.1.1)存在最优解,其中$f(x)$,$g_i(x)$ $(i=1,2,\cdots,m)$和$h_j(x)$ $(j=1,2,\cdots,l)$是$\mathbb{R}^n$上的实值连续函数.又设$\{\sigma_k\}$是趋向无穷大的严格单调递增正数列,且$k=1$时,无约束问题(10.1.11)存在最优解x_k,那么

(1) 对每个k,无约束问题(10.1.11)都存在最优解;

(2) 若对某个$\bar{k}$,有$x_{\bar{k}}\in S$,则$x_{\bar{k}}$为问题(10.1.1)的最优解;

(3) 若情形(2)总不发生,则得到点列$\{x_k\}$,其任意极限点都是问题(10.1.1)的最优解.

证明 (1)和(2)已由定理10.1.1得证.

(3) 由所给条件,设x^*为问题(10.1.1)的最优解,$\bar{x}$为$\{x_k\}$的一个极限点,不妨就设$\lim\limits_{k\to\infty}x_k=\bar{x}$,则由$x^*\in S$有

$$\begin{aligned}f(x_k)&\leqslant f(x_k)+\sigma_kP(x_k)=F(x_k,\sigma_k)\\&\leqslant F(x^*,\sigma_k)=f(x^*).\end{aligned}$$

又$f(x)$为连续函数,令$k\to\infty$,可得

$$f(\bar{x}) \leqslant f(x^*). \tag{10.1.14}$$

下面证明$\bar{x} \in S$.

由引理 10.1.2 知,$\{F(x_k,\sigma_k)\}$及$\{f(x_k)\}$都是单调递增数列,又由上面的证明有

$$F(x_k,\sigma_k) \leqslant f(x^*), k=1,2,\cdots,$$

$$f(x_k) \leqslant f(x^*), k=1,2,\cdots,$$

故知$\{F(x_k,\sigma_k)\}$及$\{f(x_k)\}$都是单调递增且有上界的数列,因此$\lim\limits_{k\to\infty} F(x_k,\sigma_k)$及$\lim\limits_{k\to\infty} f(x_k)$都存在.

由于

$$P(x_k) = \frac{F(x_k,\sigma_k) - f(x_k)}{\sigma_k}, k=1,2,\cdots,$$

令 $k\to\infty$,两端同时取极限,则根据$\lim\limits_{k\to\infty}\sigma_k = \infty$ 可得

$$\lim_{k\to\infty} P(x_k) = 0.$$

再由 $g_i(x)$和 $h_j(x)$的连续性,有

$$P(\bar{x}) = \lim_{k\to\infty} P(x_k) = 0,$$

因此,按罚函数的定义可知$\bar{x} \in S$.

又 x^* 为问题(10.1.1)的最优解,故由$\bar{x} \in S$ 有

$$f(x^*) \leqslant f(\bar{x}). \tag{10.1.15}$$

综合(10.1.14)式和(10.1.15)式,即知

$$f(\bar{x}) = f(x^*),$$

这说明$\bar{x}$是问题(10.1.1)的最优解. □

由上述定理的证明过程,我们还可得到一个重要结论:

$$\lim_{k\to\infty}\sigma_k P(x_k) = 0.$$

事实上,由$f(x_k) \leqslant F(x_k,\sigma_k)$可推得

$$\lim_{k\to\infty} f(x_k) \leqslant \lim_{k\to\infty} F(x_k,\sigma_k);$$

另一方面,由于$f(x)$连续,因此

$$\lim_{k\to\infty} f(x_k) = f(\lim_{k\to\infty} x_k) = f(\bar{x}) = f(x^*).$$

从而,由 $F(x_k,\sigma_k) \leqslant f(x^*)$推得

$$\lim_{k\to\infty} F(x_k,\sigma_k) \leqslant \lim_{k\to\infty} f(x_k).$$

于是

$$\lim_{k\to\infty} F(x_k,\sigma_k) = \lim_{k\to\infty} f(x_k).$$

所以

$$\lim_{k\to\infty}\sigma_k P(x_k) = \lim_{k\to\infty}[F(x_k,\sigma_k) - f(x_k)] = 0.$$

这就是取 $\sigma_k P(\boldsymbol{x}_k) < \varepsilon$ 作为迭代终止准则的原因.

10.2 内罚函数法

在外罚函数法中,$\{\boldsymbol{x}_k\}$ 是从可行域 S 的外部逼近最优解 $\bar{\boldsymbol{x}}$ 的. 当我们在某个充分大的 σ_k 处终止迭代时,近似最优解 $\boldsymbol{x}_k$ 一般只近似地满足约束条件,对于某些实际问题,这样的解是不能被接受的. 为了使迭代点总是可行的,可以采用本节将要介绍的内罚函数法(internal penalty function method).

内罚函数法在迭代中总是从可行点出发,并保持在可行域内部进行搜索. 因此,这种方法适用于只有不等式约束的最优化问题(10.1.4).

对于不等式约束问题(10.1.4),其可行域 S 的内部

$$\text{int}S = \{\boldsymbol{x} \in \mathbb{R}^n \mid g_i(\boldsymbol{x}) > 0, i = 1,2,\cdots,m\}.$$

为了保持迭代点始终含于 $\text{int}S$,我们引入另一种罚函数

$$G(\boldsymbol{x},r) = f(\boldsymbol{x}) + rB(\boldsymbol{x}),$$

其中参数 r 是很小的正数,$B(\boldsymbol{x})$ 是 $\text{int}S$ 上的非负实值连续函数,当点 $\boldsymbol{x}$ 趋向可行域 S 的边界时,$B(\boldsymbol{x}) \to \infty$.

显然,罚函数 $G(\boldsymbol{x},r)$ 的作用是对企图脱离可行域的点给予惩罚,相当于在可行域的边界设置了障碍,不让迭代点穿越到可行域之外,因此也称为障碍函数(barrier function).

两种常用的形式为

$$G(\boldsymbol{x},r) = f(\boldsymbol{x}) + r\sum_{i=1}^{m}\frac{1}{g_i(\boldsymbol{x})},$$

及

$$G(\boldsymbol{x},r) = f(\boldsymbol{x}) - r\sum_{i=1}^{m}\ln g_i(\boldsymbol{x}),$$

分别称为倒数障碍函数和对数障碍函数.

根据 $G(\boldsymbol{x},r)$ 的定义可知,r 取值越小,问题

$$\begin{cases}\min & G(\boldsymbol{x},r) = f(\boldsymbol{x}) + rB(\boldsymbol{x});\\ \text{s.t.} & \boldsymbol{x} \in \text{int}S\end{cases} \tag{10.2.1}$$

的最优解越接近问题(10.1.4)的最优解. 但是,这里存在与外罚函数法类似的问题,即如果 r 太小,则会给问题(10.2.1)的求解带来很大困难. 因此,我们仍利用序列无约束极小化方法(SUMT),取一个严格单调递减趋于 0 的正数列 $\{r_k\}$,对每个 k,从可行域的内部 $\text{int}S$ 出发,求解问题

$$\begin{cases}\min & G(\boldsymbol{x},r_k) = f(\boldsymbol{x}) + r_kB(\boldsymbol{x});\\ \text{s.t.} & \boldsymbol{x} \in \text{int}S.\end{cases} \tag{10.2.2}$$

若问题(10.1.4)的最优解含于 int$S$,则当 r_k 取到一适当的值时,问题(10.2.2)的最优解可以达到它;若问题(10.1.4)的最优解落在 S 的边界上,则随着 r_k 的减小,问题(10.2.2)的最优解点列将向 S 边界上的该最优解逼近. 这种利用罚函数生成一系列内点逼近原约束问题最优解的方法称为内罚函数法或 SUMT 内点法.

算法 10-2(内罚函数法)

Step 1　选取初始数据. 给定初始内点 $\boldsymbol{x}_0 \in \text{int}S$,初始参数 $r_1 > 0$,缩小系数 $\beta \in (0,1)$,允许误差 $\varepsilon > 0$,令 $k=1$.

Step 2　求解无约束问题. 以 $\boldsymbol{x}_{k-1}$ 为初始点,求解无约束问题(10.2.2),设其最优解为 $\boldsymbol{x}_k$.

Step 3　检查是否满足终止准则. 若 $r_k B(\boldsymbol{x}_k) < \varepsilon$,则迭代终止,$\boldsymbol{x}_k$ 为约束问题(10.1.4)的近似最优解;否则,令 $r_{k+1} = \beta r_k, k := k+1$,返回 Step 2.

在上述内罚函数法中,问题(10.2.2)实际上仍是约束问题,且约束条件看上去比问题(10.1.4)的约束条件更复杂. 但是,只要初始迭代点从可行域的内部选取,$B(\boldsymbol{x})$ 的障碍作用会自动实现,保证问题(10.2.2)的最优解仍会含于 int$S$. 因此,从计算的角度来看,问题(10.2.2)可当做无约束问题来处理,仅有的差别是在求极小过程中进行一维搜索时要适当控制步长,以免迭代点脱离可行域而导致迭代提前结束.

例 10.2.1　利用对数障碍函数求解不等式约束问题

$$\begin{cases} \min & f(\boldsymbol{x}) = x_1 + 2x_2; \\ \text{s.t.} & -x_1^2 + x_2 \geqslant 0, \\ & x_1 \geqslant 0, \end{cases} \tag{10.2.3}$$

要求选取 $r_1 = 4, \beta = \dfrac{1}{2}$.

解　引入对数障碍函数

$$G(\boldsymbol{x}, r_k) = (x_1 + 2x_2) - r_k[\ln(-x_1^2 + x_2) + \ln x_1],$$

我们用解析方法求解问题

$$\begin{cases} \min & G(\boldsymbol{x}, r_k); \\ \text{s.t.} & \boldsymbol{x} \in \text{int}S, \end{cases} \tag{10.2.4}$$

其中 $r_1 = 4, r_{k+1} = \dfrac{1}{2} r_k, k = 1,2,\cdots$. 令

$$\begin{cases} \dfrac{\partial G}{\partial x_1} = 1 + \dfrac{2x_1 r_k}{-x_1^2 + x_2} - \dfrac{r_k}{x_1} = 0, \\ \dfrac{\partial G}{\partial x_2} = 2 - \dfrac{r_k}{-x_1^2 + x_2} = 0, \end{cases}$$

解得 $\boldsymbol{x}_k=(x_1^{(k)},x_2^{(k)})^{\mathrm{T}}$，其中

$$x_1^{(k)}=\frac{1}{8}(-1+\sqrt{16r_k+1}),x_2^{(k)}=\frac{1}{64}(-1+\sqrt{16r_k+1})^2+\frac{1}{2}r_k.$$

依次对 $k=1,2,\cdots$ 用上述公式计算，结果如表 10.2.1 所示.

表 10.2.1　r 值对应的解

k	r_k	$\boldsymbol{x}_k$	k	r_k	$\boldsymbol{x}_k$
1	4	$(0.885,2.781)^{\mathrm{T}}$	7	0.063	$(0.057,0.035)^{\mathrm{T}}$
2	2	$(0.593,1.350)^{\mathrm{T}}$	8	0.032	$(0.029,0.017)^{\mathrm{T}}$
3	1	$(0.391,0.653)^{\mathrm{T}}$	9	0.016	$(0.016,0.008)^{\mathrm{T}}$
4	0.5	$(0.375,0.390)^{\mathrm{T}}$	10	0.008	$(0.008,0.004)^{\mathrm{T}}$
5	0.25	$(0.155,0.149)^{\mathrm{T}}$	11	0.004	$(0.003,0.002)^{\mathrm{T}}$
6	0.125	$(0.090,0.071)^{\mathrm{T}}$	12	0.002	$(0.002,0.001)^{\mathrm{T}}$

表中每一个值都对应 int S 中的一个可行内点，如图 10.2.1 所示，其中 $f=x_1+2x_2$. 从图 10.2.1 中可以看出，当 $k\to\infty$ 即 $r_k\to 0$ 时，无约束问题(10.2.4)的最优解点列 $\{\boldsymbol{x}_k\}$ 从约束问题(10.2.3)的可行域内部

$\mathrm{int}S=\{(x_1,x_2)^{\mathrm{T}}\mid -x_1^2+x_2>0,x_1>0\}$，

向可行域边界上的最优解 $\bar{\boldsymbol{x}}=(0,0)^{\mathrm{T}}$ 逼近. □

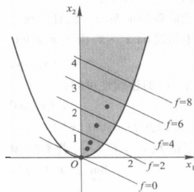


图 10.2.1　例 10.2.1 解的迭代进程

内罚函数法的收敛性由下面的定理保证.

定理 10.2.1　设不等式约束问题(10.1.4)的可行域内部 int$S$ 非空且存在最优解，其中 $f(\boldsymbol{x})$ 和 $g_i(\boldsymbol{x})(i=1,2,\cdots,m)$ 是 $\mathbb{R}^n$ 上的实值连续函数. 又设 $\{r_k\}$ 是严格单调递减趋于 0 的正数列，且对每个 k，无约束问题(10.2.2)存在最优解 $\boldsymbol{x}_k$，则点列 $\{\boldsymbol{x}_k\}$ 的任意极限点都是问题(10.1.4)的最优解.

证明　先证 $\{G(\boldsymbol{x}_k,r_k)\}$ 是单调递减且有下界的数列. 设 $\boldsymbol{x}^*$ 为问题(10.1.4)的最优解，由于 $\boldsymbol{x}_k\in\mathrm{int}\,S$，则

$$\begin{aligned}G(\boldsymbol{x}_{k+1},r_{k+1})&\leqslant G(\boldsymbol{x}_k,r_{k+1})=f(\boldsymbol{x}_k)+r_{k+1}B(\boldsymbol{x}_k)\\&\leqslant f(\boldsymbol{x}_k)+r_kB(\boldsymbol{x}_k)=G(\boldsymbol{x}_k,r_k),\end{aligned}$$

且

$$G(\boldsymbol{x}_k,r_k)=f(\boldsymbol{x}_k)+r_kB(\boldsymbol{x}_k)\geqslant f(\boldsymbol{x}_k)\geqslant f(\boldsymbol{x}^*),$$

即$\{G(\boldsymbol{x}_k,r_k)\}$单调递减且有下界. 因此存在极限(记为$G_0$)且$G_0\geqslant f(\boldsymbol{x}^*)$. 下面证明$G_0=f(\boldsymbol{x}^*)$.

用反证法,假若$G_0>f(\boldsymbol{x}^*)$,则必存在$\varepsilon>0$,使得

$$G_0>f(\boldsymbol{x}^*)+2\varepsilon. \tag{10.2.5}$$

由$f(\boldsymbol{x})$的连续性可知,存在可行点$\hat{\boldsymbol{x}}\in S$,使

$$f(\hat{\boldsymbol{x}})<f(\boldsymbol{x}^*)+\varepsilon.$$

因为

$$G(\boldsymbol{x}_k,r_k)\leqslant G(\hat{\boldsymbol{x}},r_k)=f(\hat{\boldsymbol{x}})+r_kB(\hat{\boldsymbol{x}})<f(\boldsymbol{x}^*)+\varepsilon+r_kB(\hat{\boldsymbol{x}}),$$

又当k充分大时,总有$r_kB(\hat{\boldsymbol{x}})<\varepsilon$,于是有

$$\lim_{k\to\infty}G(\boldsymbol{x}_k,r_k)\leqslant f(\boldsymbol{x}^*)+\varepsilon+\varepsilon,$$

即

$$G_0\leqslant f(\boldsymbol{x}^*)+2\varepsilon,$$

此与(10.2.5)式矛盾.

然后我们证明$\{\boldsymbol{x}_k\}$的任意极限点$\bar{\boldsymbol{x}}$都是问题(10.1.4)的最优解.

不妨设$\lim\limits_{k\to\infty}\boldsymbol{x}_k=\bar{\boldsymbol{x}}$. 由于$\boldsymbol{x}_k\in \text{int}S$,因此有

$$g_i(\boldsymbol{x}_k)>0\quad(i=1,2,\cdots,m),$$

根据$g_i(\boldsymbol{x})$的连续性可知,

$$g_i(\bar{\boldsymbol{x}})\geqslant 0\quad(i=1,2,\cdots,m),$$

即$\bar{\boldsymbol{x}}\in S$. 因此$f(\bar{\boldsymbol{x}})\geqslant f(\boldsymbol{x}^*)$.

下证$f(\bar{\boldsymbol{x}})=f(\boldsymbol{x}^*)$. 若不然,则$f(\bar{\boldsymbol{x}})>f(\boldsymbol{x}^*)$且

$$\lim_{k\to\infty}[f(\boldsymbol{x}_k)-f(\boldsymbol{x}^*)]=f(\bar{\boldsymbol{x}})-f(\boldsymbol{x}^*)>0,$$

又

$$G(\boldsymbol{x}_k,r_k)-f(\boldsymbol{x}^*)=f(\boldsymbol{x}_k)-f(\boldsymbol{x}^*)+r_kB(\boldsymbol{x}_k)\geqslant f(\boldsymbol{x}_k)-f(\boldsymbol{x}^*),$$

故

$$\lim_{k\to\infty}[G(\boldsymbol{x}_k,r_k)-f(\boldsymbol{x}^*)]\neq 0,$$

即$G_0\neq f(\boldsymbol{x}^*)$,此与上面证明的$G_0=f(\boldsymbol{x}^*)$矛盾. 故$f(\bar{\boldsymbol{x}})=f(\boldsymbol{x}^*)$,即$\bar{\boldsymbol{x}}$是问题(10.1.4)的最优解. □

在内罚函数法中必须先知道一个初始内点$\boldsymbol{x}_0\in\text{int}S$. 在实际问题中,如果初始内点不能凭直观找出时,则必须给出一种求初始内点的算法. 下面介绍一个寻找初始内点的迭代算法,它是基于内罚函数法的思想而得到的.

算法 10-3(求内罚函数法中初始内点的算法)

Step 1 选取初始数据. 给定初始点$\boldsymbol{x}_0$,初始参数$r_1>0$,缩小系数$\beta\in(0,$

1)，令 $k=0$.

Step 2　确定指标集. 令

$$I_k=\{i\mid g_i(\boldsymbol{x}_k)\leqslant 0, i=1,2,\cdots,m\},$$
$$J_k=\{i\mid g_i(\boldsymbol{x}_k)>0, i=1,2,\cdots,m\}.$$

Step 3　检验是否满足终止准则. 若 $I_k=\varnothing$，则 $\boldsymbol{x}_k\in \text{int}S$，迭代终止；否则，转 Step 4.

Step 4　求解无约束问题. 以 $\boldsymbol{x}_k$ 为初始点，求无约束问题 $\min\limits_{x\in S_k} H_{k+1}(\boldsymbol{x})$ 的最优解 $\boldsymbol{x}_{k+1}$，其中

$$H_{k+1}(\boldsymbol{x})=-\sum_{i\in I_k} g_i(\boldsymbol{x})+r_{k+1}\sum_{i\in J_k}\frac{1}{g_i(\boldsymbol{x})},$$
$$S_k=\{\boldsymbol{x}\mid g_i(\boldsymbol{x})>0, i\in J_k\}.$$

令 $r_{k+2}=\beta r_{k+1}$，$k:=k+1$，返回 Step 2.

不难看出，在算法 10－3 中，对于 $k=0,1,2,\cdots$，有

$$I_{k+1}\subseteq I_k, J_{k+1}\supseteq J_k,$$

并且还有

定理 10.2.2　假设在算法 10－3 中，$\text{int}S\neq\varnothing$，$\{r_k\}$ 为严格单调递减趋于 0 的正数列，并且问题 $\min\limits_{x\in S_{k-1}} H_k(\boldsymbol{x})$ 的最优解为 $\boldsymbol{x}_k$，则该算法必在有限步结束.

证明　反证法. 若不然，则由

$$\{1,2,\cdots,m\}\supseteq I_0\supseteq I_1\supseteq\cdots\supseteq I_k\supseteq I_{k+1}\supseteq\cdots$$

可知，必存在 $k_1\geqslant 0$，使得

$$I=I_{k_1}=I_{k_1+1}=I_{k_1+2}=\cdots, I\neq\varnothing.$$

取定 $\boldsymbol{y}_0\in\text{int}S$，则当 $k\geqslant k_1$ 时，有

$$-\sum_{i\in I} g_i(\boldsymbol{x}_k)+r_k\sum_{i\in J}\frac{1}{g_i(\boldsymbol{x}_k)}\leqslant-\sum_{i\in I} g_i(\boldsymbol{y}_0)+r_k\sum_{i\in J}\frac{1}{g_i(\boldsymbol{y}_0)}, \tag{10.2.6}$$

其中 $J=\{1,2,\cdots,m\}\setminus I$. 由于 $\lim\limits_{k\to\infty} r_k=0$，$I\neq\varnothing$，$\boldsymbol{y}_0\in\text{int}S$，因此存在 $k_2>k_1$，使得

$$-\sum_{i\in I} g_i(\boldsymbol{y}_0)+r_{k_2}\sum_{i\in J}\frac{1}{g_i(\boldsymbol{y}_0)}<0. \tag{10.2.7}$$

由 $k_2>k_1$ 可知 $I_{k_2}=I$，从而由(10.2.6)和(10.2.7)两式得

$$-\sum_{i\in I_{k_2}} g_i(\boldsymbol{x}_{k_2})+r_{k_2}\sum_{i\in J_{k_2}}\frac{1}{g_i(\boldsymbol{x}_{k_2})}<0.$$

另一方面，由 I_{k_2} 和 J_{k_2} 的定义显然有

$$-\sum_{i\in I_{k_2}} g_i(\boldsymbol{x}_{k_2})+r_{k_2}\sum_{i\in J_{k_2}}\frac{1}{g_i(\boldsymbol{x}_{k_2})}\geqslant 0,$$

此为矛盾. □

在本节的开始我们曾指出，内罚函数法只能求解不等式约束问题. 为了利用这个算法去求解一般的约束非线性最优化问题，可以把外罚函数法和内罚函数法结合起来，形成所谓的混合罚函数法，其适用范围更广（参见文献[19,20]).

上面介绍的外罚函数法和内罚函数法均采用序列无约束极小化技巧，结构简单，使用方便，并能求解导数不存在的问题，因此得到了广泛的应用. 但由于要求解一系列无约束问题，工作量必然增大，且随着罚因子趋向其极限，罚函数的Hesse 矩阵的条件数无限增大，变得越来越病态，这给无约束极小化的数值计算带来了很大的困难，甚至由于舍入误差的影响，使求得的 $\boldsymbol{x}_k$ 失真.

10.3 广义乘子法

为了利用外罚函数法的思想并克服它的缺点，本节考虑把罚函数与 Lagrange 函数结合起来，构造出更合适的新目标函数，使得在罚因子 σ 适当大的情况下，借助于 Lagrange 乘子就能逐步达到原约束问题的最优解. 由于这种方法要借助于 Lagrange 乘子的迭代进行求解而又区别于经典的 Lagrange 乘子法，故称为广义乘子法.

10.3.1 等式约束下的广义乘子法

现在我们考虑只带等式约束的最优化问题(10.1.2). 不难知道，问题(10.1.2)等价于问题

$$\begin{cases} \min \quad \left[f(\boldsymbol{x}) + \dfrac{\sigma}{2}\sum_{j=1}^{l} h_j^2(\boldsymbol{x})\right]; \\ \text{s.t.} \quad h_j(\boldsymbol{x}) = 0, j = 1,2,\cdots,l, \end{cases} \tag{10.3.1}$$

其中 $\sigma > 0$. 该问题的 Lagrange 函数

$$L(\boldsymbol{x},\sigma,\boldsymbol{v}) = f(\boldsymbol{x}) + \frac{\sigma}{2}\sum_{j=1}^{l} h_j^2(\boldsymbol{x}) - \sum_{j=1}^{l} v_j h_j(\boldsymbol{x}),$$

由于 $L(\boldsymbol{x},\sigma,\boldsymbol{v})$ 中既有罚项 $\dfrac{\sigma}{2}\sum_{j=1}^{l} h_j^2(\boldsymbol{x})$，又有乘子项 $-\sum_{j=1}^{l} v_j h_j(\boldsymbol{x})$，故称为乘子罚函数(multiplier penalty function).

与外罚函数类似，若设 $\{\sigma_k\}$ 为单调递增的正数列，则我们可以把求解等式约束问题(10.1.2)转化为求解一系列的无约束问题

$$\min_{\boldsymbol{x}\in \mathbf{R}^n} L(\boldsymbol{x},\sigma_k,\boldsymbol{v}_k) = f(\boldsymbol{x}) + \frac{\sigma_k}{2}\sum_{j=1}^{l} h_j^2(\boldsymbol{x}) - \sum_{j=1}^{l} v_j^{(k)} h_j(\boldsymbol{x}), \tag{10.3.2}$$

其中 $\boldsymbol{v}_k = (v_1^{(k)}, v_2^{(k)}, \cdots, v_l^{(k)})^{\mathrm{T}}$ 是第 k 次迭代中采用的 Lagrange 乘子，并且有

定理 10.3.1 设 $\boldsymbol{x}_k$ 是无约束问题(10.3.2)的最优解，则 $\boldsymbol{x}_k$ 也是问题

$$\begin{cases} \min \quad f(\boldsymbol{x}); \\ \text{s.t.} \quad h_j(\boldsymbol{x}) = h_j(\boldsymbol{x}_k), j=1,2,\cdots,l \end{cases} \tag{10.3.3}$$

的最优解.

证明 由于 $\boldsymbol{x}_k$ 是问题(10.3.2)的最优解,因此对于任意 $\boldsymbol{x}\in\mathbb{R}^n$,都有

$$L(\boldsymbol{x}_k,\boldsymbol{\sigma}_k,\boldsymbol{v}_k)\leqslant L(\boldsymbol{x},\boldsymbol{\sigma}_k,\boldsymbol{v}_k),$$

即

$$f(\boldsymbol{x}_k)+\frac{\sigma_k}{2}\sum_{j=1}^{l}h_j^2(\boldsymbol{x}_k)-\sum_{j=1}^{l}v_j^{(k)}h_j(\boldsymbol{x}_k)$$

$$\leqslant f(x)+\frac{\sigma_k}{2}\sum_{j=1}^{l}h_j^2(x)-\sum_{j=1}^{l}v_j^{(k)}h_j(\boldsymbol{x}),$$

也即

$$f(\boldsymbol{x}_k)-f(\boldsymbol{x})\leqslant\frac{\sigma_k}{2}\sum_{j=1}^{l}\left[h_j^2(\boldsymbol{x})-h_j^2(\boldsymbol{x}_k)\right]-$$

$$\sum_{j=1}^{l}v_j^{(k)}\left[h_j(\boldsymbol{x})-h_j(\boldsymbol{x}_k)\right].$$

当 $\boldsymbol{x}$ 满足约束条件

$$h_j(\boldsymbol{x})=h_j(\boldsymbol{x}_k), j=1,2,\cdots,l$$

时,$f(\boldsymbol{x}_k)\leqslant f(\boldsymbol{x})$. 因此,$\boldsymbol{x}_k$ 也是问题(10.3.3)的最优解. □

由定理 10.3.1 可知,经过适当调整 σ_k 与 $\boldsymbol{v}_k$,当

$$h_j(\boldsymbol{x}_k)=0, j=1,2,\cdots,l$$

成立时,$\boldsymbol{x}_k$ 即为等式约束问题(10.1.2)的最优解. 因此,在下面叙述的算法步骤中,我们将把 $\boldsymbol{x}_k$ 是否近似地满足问题(10.1.2)的约束条件,即

$$\|\boldsymbol{h}(\boldsymbol{x}_k)\|<\varepsilon$$

作为终止准则,这里 $\varepsilon>0$ 是预先给定的允许误差,向量函数 $\boldsymbol{h}(\boldsymbol{x})=(h_1(\boldsymbol{x}), h_2(\boldsymbol{x}),\cdots,h_l(\boldsymbol{x}))^{\mathrm{T}}$.

现在,算法的困难就在于 σ_k 与 $\boldsymbol{v}_k$ 的选取.

如果 $f(\boldsymbol{x})$ 和 $h_j(\boldsymbol{x})(j=1,2,\cdots,l)$ 均具有一阶连续偏导数,那么

$$\nabla_x L(\boldsymbol{x},\sigma_k,\boldsymbol{v}_k)=\nabla f(\boldsymbol{x})-\sum_{j=1}^{l}\left[v_j^{(k)}-\sigma_k h_j(\boldsymbol{x})\right]\nabla h_j(\boldsymbol{x}).$$

当给定了充分大的 σ_k,并且在 $\boldsymbol{v}_k$ 已知的情况下求得无约束问题(10.3.2)的最优解 $\boldsymbol{x}_k$ 时,必有

$$\nabla_x L(\boldsymbol{x}_k,\sigma_k,\boldsymbol{v}_k)=\nabla f(\boldsymbol{x}_k)-\sum_{j=1}^{l}\left[v_j^{(k)}-\sigma_k h_j(\boldsymbol{x}_k)\right]\nabla h_j(\boldsymbol{x}_k)=\boldsymbol{0}.$$

因此,为了使 $\boldsymbol{v}_k$ 逼近问题(10.1.2)的 Lagrange 乘子 $\boldsymbol{v}$,可采用乘子迭代

$$v_j^{(k+1)}=v_j^{(k)}-\sigma_k h_j(\boldsymbol{x}_k), j=1,2,\cdots l,$$

来修正 $\boldsymbol{v}_k$，在下一次迭代中便求问题

$$\min_{x\in\mathbb{R}^n} L(\boldsymbol{x},\sigma_k,\boldsymbol{v}_{k+1})$$

的最优解 $\boldsymbol{x}_{k+1}$. 这样操作下去，直到 $\boldsymbol{v}_k$ 收敛，即直到

$$\|\boldsymbol{h}(\boldsymbol{x}_k)\| < \varepsilon$$

为止. 如果 $h_j(\boldsymbol{x}_k)$ 不收敛到 0 或者收敛太慢，就增大 σ_k 的值，再进行迭代. 收敛快慢一般用 $\dfrac{\|\boldsymbol{h}(\boldsymbol{x}_k)\|}{\|\boldsymbol{h}(\boldsymbol{x}_{k-1})\|}$ 来衡量.

综上所述，下面给出等式约束下的广义乘子法的迭代步骤.

算法 10-4(等式约束下的广义乘子法)

Step 1　选取初始数据. 给定初始点 $\boldsymbol{x}_0$，初始乘子 $\boldsymbol{v}_1$，初始罚因子 $\sigma_1>0$，放大系数 $\alpha>1$，允许误差 $\varepsilon>0$，参数 $\gamma\in(0,1)$，令 $k=1$.

Step 2　求解无约束问题. 以 $\boldsymbol{x}_{k-1}$ 为初始点，求解无约束问题(10.3.2)，设其最优解为 $\boldsymbol{x}_k$.

Step 3　检查是否满足终止准则. 若

$$\|\boldsymbol{h}(\boldsymbol{x}_k)\| < \varepsilon,$$

则迭代终止，$\boldsymbol{x}_k$ 为等式约束问题(10.1.2)的近似最优解；否则，转 Step 4.

Step 4　判断收敛快慢. 若

$$\frac{\|\boldsymbol{h}(\boldsymbol{x}_k)\|}{\|\boldsymbol{h}(\boldsymbol{x}_{k-1})\|}\geqslant\gamma,$$

则令 $\sigma_{k+1}=\alpha\sigma_k$，转 Step 5；否则，令 $\sigma_{k+1}:=\sigma_k$，转 Step 5；

Step 5　进行乘子迭代. 令

$$v_j^{(k+1)}=v_j^{(k)}-\sigma_k h_j(\boldsymbol{x}_k),\, j=1,2,\cdots,l,$$

及 $k:=k+1$，返回 Step 2.

例 10.3.1　用等式约束下的广义乘子法求解问题

$$\begin{cases}\min \quad f(\boldsymbol{x})=\dfrac{1}{2}x_1^2+\dfrac{1}{6}x_2^2;\\[2mm] \text{s.t.} \quad x_1+x_2=1,\end{cases}$$

要求选取初始乘子 $v_1=0$，初始罚因子 $\sigma_1=0.1$，$\alpha=2$.

解　构造乘子罚函数

$$L(\boldsymbol{x},\sigma_k,v_k)=\frac{1}{2}x_1^2+\frac{1}{6}x_2^2+\frac{\sigma_k}{2}(x_1+x_2-1)^2-v_k(x_1+x_2-1),$$

其中 $\sigma_{k+1}=2\sigma_k$，$k=1,2,\cdots$.

令 $\nabla L(\boldsymbol{x},\sigma_k,v_k)=\boldsymbol{0}$，用解析方法可求得

$$\min_{x\in\mathbb{R}^n} L(\boldsymbol{x},\sigma_k,v_k)$$

的最优解

$$\boldsymbol{x}_k=\left(\frac{\sigma_k+v_k}{1+4\sigma_k},\frac{3(\sigma_k+v_k)}{1+4\sigma_k}\right)^{\mathrm{T}},k=1,2,\cdots.$$

按算法 10 - 4 的 Step 5 中的乘子迭代公式,有

$$v_{k+1}=v_k-\sigma_k(x_1^{(k)}+x_2^{(k)}-1),k=1,2,\cdots,$$

其中 $\boldsymbol{x}_k=(x_1^{(k)},x_2^{(k)})^{\mathrm{T}},v_1=0,\sigma_1=0.1$.

依次对 $k=1,2,\cdots$ 用上述公式,得到 σ_k,v_k 和 $\boldsymbol{x}_k$ 的数值结果,如表 10.3.1 所示. □

表 10.3.1 例 10.3.1 的迭代过程

k	σ_k	v_k	$\boldsymbol{x}_k$
1	0.1	0	$(0.071\ 4,0.214\ 2)^{\mathrm{T}}$
2	0.2	0.071 4	$(0.150\ 7,0.452\ 3)^{\mathrm{T}}$
3	0.4	0.150 7	$(0.211\ 8,0.635\ 5)^{\mathrm{T}}$
4	0.8	0.211 8	$(0.340\ 9,0.722\ 7)^{\mathrm{T}}$
5	1.6	0.240 9	$(0.248\ 7,0.746\ 3)^{\mathrm{T}}$
6	3.2	0.248 7	$(0.249\ 9,0.749\ 7)^{\mathrm{T}}$
7	6.4	0.249 9	$(0.249\ 9,0.749\ 9)^{\mathrm{T}}$

在例 10.1.2 中,我们曾用外罚函数法求解过本例的问题,在那里进行了 16 轮迭代才获得原问题的近似最优解(见表 10.1.2)

$$\boldsymbol{x}_{16}=(0.249\ 9,0.749\ 9)^{\mathrm{T}}.$$

现在用广义乘子法求解,仅 7 轮迭代就达到了同样的效果. 并且,在用外罚函数法求解时,罚因子 σ_k 要增大到 1 638.4 才能得到这一精度的近似解,但用广义乘子法,罚因子 σ_k 仅需增大到 6.4 即可.

从上述例子中我们还发现,尽管乘子迭代要增加计算量,但正是由于乘子 $\boldsymbol{v}_k$ 的逐次调整才加速了 $\boldsymbol{x}_k$ 向最优解 $\bar{\boldsymbol{x}}=(0.25,0.75)^{\mathrm{T}}$ 接近,同时也避免了因罚因子 σ_k 过大而带来的数值计算上的困难,这充分体现了广义乘子法比外罚函数法优越的方面.

10.3.2 不等式约束下的广义乘子法

下面考虑带有不等式约束的问题(10.1.1)和(10.1.4).

对于只带不等式约束的问题(10.1.4),为了利用关于等式约束问题所得到的结论,引入附加变量 $\boldsymbol{y}=(y_1,y_2,\cdots,y_m)^{\mathrm{T}}$,则问题(10.1.4)等价于等式约束问题

$$\begin{cases}\min & f(\boldsymbol{x});\\ \text{s.t} & g_i(\boldsymbol{x})-y_i^2=0,i=1,2,\cdots,m.\end{cases}\tag{10.3.4}$$

问题(10.3.4)所对应的广义乘子法中的乘子罚函数为

$$\bar{L}(\boldsymbol{x},\boldsymbol{y},\sigma_k,\boldsymbol{w}_k)=f(\boldsymbol{x})+\frac{\sigma_k}{2}\sum_{i=1}^{m}[g_i(\boldsymbol{x})-y_i^2]^2-\sum_{i=1}^{m}w_i^{(k)}[g_i(\boldsymbol{x})-y_i^2],$$

从而把问题(10.1.4)转化为求解无约束问题

$$\min_{x\in\mathbb{R}^n,y\in\mathbb{R}^m}\bar{L}(\boldsymbol{x},\boldsymbol{y},\sigma_k,\boldsymbol{w}_k).\qquad(10.3.5)$$

为了把问题(10.3.5)化成只关于 $\boldsymbol{x}$ 求极小的问题,我们将先对函数 $\bar{L}(\boldsymbol{x},\boldsymbol{y},\sigma_k,\boldsymbol{w}_k)$ 关于 $\boldsymbol{y}$ 求极小,由此解出 $\boldsymbol{y}$,再代入问题(10.3.5)中.

用配方法将 $\bar{L}(\boldsymbol{x},\boldsymbol{y},\sigma_k,\boldsymbol{w}_k)$ 转化为

$$\begin{aligned}&\bar{L}(\boldsymbol{x},\boldsymbol{y},\sigma_k,\boldsymbol{w}_k)\\&=f(\boldsymbol{x})+\sum_{i=1}^{m}\left\{\frac{\sigma_k}{2}\left[y_i^2-\frac{1}{\sigma_k}(\sigma_k g_i(\boldsymbol{x})-w_i^{(k)})\right]^2-\frac{(w_i^{(k)})^2}{2\sigma_k}\right\},\end{aligned}\qquad(10.3.6)$$

为使 $\bar{L}$ 关于 y_i 取极小值,则 y_i 应满足

(1) 当 $\sigma_k g_i(\boldsymbol{x})-w_i^{(k)}\geqslant 0$ 时,$y_i^2=\frac{1}{\sigma_k}[\sigma_k g_i(\boldsymbol{x})-w_i^{(k)}]$;

(2) 当 $\sigma_k g_i(\boldsymbol{x})-w_i^{(k)}<0$ 时,$y_i=0$.

综合(1)、(2)即有

$$y_i^2=\frac{1}{\sigma_k}\max\{0,\sigma_k g_i(\boldsymbol{x})-w_i^{(k)}\},i=1,2,\cdots,m.\qquad(10.3.7)$$

把(10.3.7)式代入(10.3.6)式中,由此定义出与 y_i 无关的乘子罚函数

$$L(\boldsymbol{x},\sigma_k,\boldsymbol{w}_k)=f(\boldsymbol{x})+\frac{1}{2\sigma_k}\sum_{i=1}^{m}\{[\max\{0,w_i^{(k)}-\sigma_k g_i(\boldsymbol{x})\}]^2-(w_i^{(k)})^2\},$$

根据等式约束下的广义乘子法的乘子迭代公式,可以推出不等式约束下的广义乘子法的乘子迭代公式:

$$\begin{aligned}w_i^{(k+1)}&=w_i^{(k)}-\sigma_k[g_i(\boldsymbol{x}_k)-y_i^2]\\&=w_i^{(k)}-\sigma_k g_i(\boldsymbol{x}_k)+\max\{0,\sigma_k g_i(\boldsymbol{x}_k)-w_i^{(k)}\}\\&=\max\{0,w_i^{(k)}-\sigma_k g_i(\boldsymbol{x}_k)\},i=1,2,\cdots,m.\end{aligned}$$

从而,问题(10.1.4)转化为求解一系列无约束问题

$$\min_{x\in\mathbb{R}^n}L(\boldsymbol{x},\sigma_k,\boldsymbol{w}_k).$$

对于既含有不等式约束又含有等式约束的一般非线性最优化问题(10.1.1),可定义乘子罚函数

$$\begin{aligned}&L(\boldsymbol{x},\sigma_k,\boldsymbol{w}_k,\boldsymbol{v}_k)\\&=f(\boldsymbol{x})+\frac{1}{2\sigma_k}\sum_{i=1}^{m}\{[\max(0,w_i^{(k)}-\sigma_k g_i(\boldsymbol{x}))]^2-(w_i^{(k)})^2\}+\\&\quad\frac{\sigma_k}{2}\sum_{j=1}^{l}h_j^2(\boldsymbol{x})-\sum_{j=1}^{l}v_j^{(k)}h_j(\boldsymbol{x}),\end{aligned}$$

其中乘子迭代公式为

$$w_i^{(k+1)}=\max\{0,w_i^{(k)}-\sigma_k g_i(\boldsymbol{x}_k)\},i=1,2,\cdots,m,$$

$$v_j^{(k+1)}=v_j^{(k)}-\sigma_k h_j(\boldsymbol{x}_k),j=1,2,\cdots,l.$$

因此,问题(10.1.1)转化为求解一系列无约束问题

$$\min_{x\in\mathbb{R}^n}L(\boldsymbol{x},\sigma_k,\boldsymbol{w}_k,\boldsymbol{v}_k).$$

上述两类问题的具体计算步骤与等式约束的情形类似.

例 10.3.2 用广义乘子法求解

$$\begin{cases}\min & x_1^2+2x_2^2;\\ \text{s. t.} & x_1+x_2\geqslant 1.\end{cases}\tag{10.3.8}$$

解 引入乘子罚函数

$$\begin{aligned}&L(x,\sigma,w_k)\\&=x_1^2+2x_2^2+\frac{1}{2\sigma}\{[\max(0,w_k-\sigma(x_1+x_2-1))]^2-w_k^2\}\\&=\begin{cases}x_1^2+2x_2^2+\dfrac{1}{2\sigma}\{[w_k-\sigma(x_1+x_2-1)]^2-w_k^2\}, & x_1+x_2-1\leqslant\dfrac{w_k}{\sigma},\\[2ex] x_1^2+2x_2^2-\dfrac{w_k^2}{2\sigma}, & x_1+x_2-1>\dfrac{w_k}{\sigma},\end{cases}\end{aligned}$$

则有

$$\frac{\partial L}{\partial x_1}=\begin{cases}2x_1-[w_k-\sigma(x_1+x_2-1)], & x_1+x_2-1\leqslant\dfrac{w_k}{\sigma},\\[2ex] 2x_1, & x_1+x_2-1>\dfrac{w_k}{\sigma},\end{cases}$$

$$\frac{\partial L}{\partial x_2}=\begin{cases}4x_2-[w_k-\sigma(x_1+x_2-1)], & x_1+x_2-1\leqslant\dfrac{w_k}{\sigma},\\[2ex] 4x_2, & x_1+x_2-1>\dfrac{w_k}{\sigma}.\end{cases}$$

令

$$\nabla L(\boldsymbol{x},\sigma,w_k)=\boldsymbol{0},$$

得到无约束问题

$$\min_{x\in\mathbb{R}^n}L(\boldsymbol{x},\sigma,w_k)\tag{10.3.9}$$

的最优解 $\boldsymbol{x}_k=(x_1^{(k)},x_2^{(k)})^{\mathrm{T}}$,其中

$$x_1^{(k)}=\frac{2(w_k+\sigma)}{4+3\sigma},x_2^{(k)}=\frac{w_k+\sigma}{4+3\sigma}.$$

取 $\sigma=2,w_1=1$,有

$$x_1=(x_1^{(1)},x_2^{(1)})^{\mathrm T}=\left(\frac{3}{5},\frac{3}{10}\right)^{\mathrm T}.$$

再修正乘子 w_1，令

$$w_2=\max\{0,w_1-\sigma(x_1^{(1)}+x_2^{(1)}-1)\}=\frac{6}{5},$$

得到

$$x_2=(x_1^{(2)},x_2^{(2)})^{\mathrm T}=\left(\frac{16}{25},\frac{8}{25}\right)^{\mathrm T}.$$

以此类推，设在第 k 次迭代中取乘子 w_k，则得到无约束问题(10.3.9)的最优解

$$x_k=(x_1^{(k)},x_2^{(k)})^{\mathrm T}=\left(\frac{1}{5}(2+w_k),\frac{1}{10}(2+w_k)\right)^{\mathrm T}.$$

再修正 w_k，有

$$w_{k+1}=\max\{0,w_k-2(x_1^{(k)}+x_2^{(k)}-1)\}=\frac{1}{5}(2w_k+4).$$

显然，按上式迭代得到的乘子序列 $\{w_k\}$ 是收敛的，且当 $k\to\infty$ 时，$w_k\to\frac{4}{3}$，$x_k\to\left(\frac{2}{3},\frac{1}{3}\right)^{\mathrm T}=\bar{x}$，其中 $\bar{x}$ 恰为原不等式约束问题(10.3.8)的最优解. □

在上述两种情形的广义乘子法中，由于罚因子 σ_k 不必趋向无穷大就能求得原约束问题的最优解，因此，避免了外罚函数法中的病态出现. 计算经验表明，广义乘子法优于外罚函数法.

习　题　十

1. 试用外罚函数法求解下列约束非线性最优化问题，要求选取 $\sigma_1=0.5,\alpha=2,\varepsilon=10^{-4}$：

(1) $\begin{cases}\min & x_1^2+x_2^2;\\ \text{s.t.} & x_1+x_2-1=0,\\ & x_1\geqslant x_2^2.\end{cases}$　(2) $\begin{cases}\min\ x_1^2+2x_2^2;\\ \text{s.t.}\ 2x_1+x_2\leqslant 2,\\ \quad\ x_1\geqslant 1.\end{cases}$

(3) $\begin{cases}\max\ x_1x_2x_3;\\ \text{s.t.}\ 72-x_1-2x_2-2x_3=0.\end{cases}$

2. 用外罚函数法求解问题

$$\begin{cases}\min & (x-1)^2;\\ \text{s.t.} & x-2\geqslant 0.\end{cases}$$

(1) 写出 $\sigma_k=0,1,10$ 时相应的罚函数并画出图形；

(2) 取 $\sigma_k=k-1(k=1,2,\cdots)$，求出近似最优解的迭代点列；

(3) 利用(2)求问题的最优解.

3. 设计一个相应的罚函数，当 $k\to1$ 时，收敛且能用来求解例 10.1.2.

4. 用内罚函数法求解下列不等式约束问题：

(1) $\begin{cases}\min & (x+1)^2;\\ \text{s.t.} & x\geqslant 0.\end{cases}$ (2) $\begin{cases}\max & x_1x_2;\\ \text{s.t.} & -2x_1+x_2+3\geqslant 0.\end{cases}$

(3) $\begin{cases}\min & x_1^2+2x_2^2;\\ \text{s.t.} & x_1+x_2-1\geqslant 0.\end{cases}$

5. 利用对数障碍函数求解问题

$$\begin{cases}\min & x_1+2x_2;\\ \text{s.t.} & -x_1^2+x_2\geqslant 0,\\ & x_1\geqslant 0.\end{cases}$$

6. 设 $G(\boldsymbol{x},r)$ 是问题(10.1.4)对应的障碍函数,若其可行域内部 $\text{int}S=\{\boldsymbol{x}\in\mathbb{R}^n \mid g_i(\boldsymbol{x})>0, i=1,2,\cdots,m\}$ 有界,试证:

$$\begin{cases}\min & G(\boldsymbol{x},r);\\ \text{s.t.} & \boldsymbol{x}\in\text{int}S\end{cases}$$

必在 $\text{int}S$ 上达到最优解.

7. 试用广义乘子法求解下列等式约束问题

(1) $\begin{cases}\min & x_1^2+x_1x_2+x_2^2-2x_2;\\ \text{s.t.} & x_1+x_2-2=0.\end{cases}$ (取 $v_1=0$)

(2) $\begin{cases}\min & x_1^2+2x_2^2;\\ \text{s.t.} & x_1+x_2-1=0.\end{cases}$ (取 $v_1=0$)

(3) $\begin{cases}\min & x_2^2-3x_1;\\ \text{s.t.} & x_1+x_2=1,\ (\text{取 } \boldsymbol{v}_1=(0,0)^{\mathrm{T}})\\ & x_1-x_2=0.\end{cases}$

要求取 $\sigma_k=2^k(k=1,2,\cdots)$,并作3轮迭代.

8. 对于约束最优化问题

$$\begin{cases}\min & x_1^3+x_2^3;\\ \text{s.t.} & x_1+x_2=1,\end{cases}$$

分别写出由外罚函数法和广义乘子法建立的最优解的迭代点列.

9. 试用广义乘子法求解下列带不等式约束的最优化问题:

(1) $\begin{cases}\min & 2x_1+3x_2;\\ \text{s.t.} & 1-2x_1^2-x_2^2\geqslant 0.\end{cases}$ (取 $w_1=0$)

(2) $\begin{cases}\min & x_1+\frac{1}{3}(x_2+1)^2;\\ \text{s.t.} & x_1\geqslant 0,\\ & x_2\geqslant 1.\end{cases}$ (取 $w_1=(0,0)^{\mathrm{T}}$)

(3) $\begin{cases}\min & x_2^2-x_1;\\ \text{s.t.} & x_1=0,\quad (\text{取 } v_1=0, w_1=1)\\ & x_2+1\geqslant 0.\end{cases}$

要求取 $\sigma_1=0.1,\sigma_{k+1}=2\sigma_k(k=1,2,\cdots)$,并进行3轮迭代.

10. 对于等式约束问题(10.1.2),在其乘子罚函数 $L(\boldsymbol{x},\sigma,\boldsymbol{v})$ 中令 $\boldsymbol{v}=\boldsymbol{0}$,即得到罚函数法中的一般罚函数;若在 $L(\boldsymbol{x},\sigma,\boldsymbol{v})$ 中令 $\boldsymbol{v}$ 取非零值并且 $\boldsymbol{v}$ 可以随迭代而不同,但不用迭代公式修正,即得到更广的外罚函数法.试证:若乘子序列 $\{\boldsymbol{v}_k\}$ 有界,则定理10.1.3对乘子罚函数 $L(\boldsymbol{x},\sigma_k,\boldsymbol{v}_k)$ 也成立.

第十一章　二次规划与凸规划

二次规划是指目标函数为二次函数、约束条件是线性等式或线性不等式的约束优化问题. 在二次规划问题中, 目标函数的 Hesse 矩阵和约束条件的 Jacobi 矩阵都是常量, 因此最优性条件和求解方法都有其特殊性. 本章首先讨论介绍二次规划的几种典型求解方法, 然后介绍求解凸规划问题的割平面法.

11.1　等式约束二次规划问题

等式约束二次规划问题可表示为

$$\begin{cases} \min & f(\boldsymbol{x}) = \dfrac{1}{2}\boldsymbol{x}^{\mathrm{T}}\boldsymbol{Q}\boldsymbol{x} + \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}; \\ \text{s.t.} & \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}, \end{cases} \tag{11.1.1}$$

其中 $\boldsymbol{Q} \in \mathbb{R}^{n\times n}$ 为对称矩阵, $\boldsymbol{A} \in \mathbb{R}^{m\times n}$ 且不妨设 rank $\boldsymbol{A} = m < n$, $\boldsymbol{c} \in \mathbb{R}^n$, $\boldsymbol{b} \in \mathbb{R}^m$.

11.1.1　Lagrange 乘子法

在第三章已经讨论过等式约束最优化问题的最优性条件, 下面针对二次规划问题的特殊性, 给出等式约束二次规划的最优性条件.

根据定理 3.2.1 可得问题(11.1.1) 局部极小点的一阶必要条件.

设 $\bar{\boldsymbol{x}}$ 为问题(11.1.1)的局部极小点, 则存在 Lagrange 乘子向量 $\bar{\boldsymbol{v}} \in \mathbb{R}^m$, 使得

$$\begin{cases} \boldsymbol{Q}\bar{\boldsymbol{x}} + \boldsymbol{c} - \boldsymbol{A}^{\mathrm{T}}\bar{\boldsymbol{v}} = \boldsymbol{0}, \\ \boldsymbol{A}\bar{\boldsymbol{x}} = \boldsymbol{b}, \end{cases} \tag{11.1.2}$$

其等价的矩阵形式为

$$\begin{bmatrix} \boldsymbol{Q} & -\boldsymbol{A}^{\mathrm{T}} \\ -\boldsymbol{A} & \boldsymbol{0} \end{bmatrix}\begin{bmatrix} \bar{\boldsymbol{x}} \\ \bar{\boldsymbol{v}} \end{bmatrix} = \begin{bmatrix} -\boldsymbol{c} \\ -\boldsymbol{b} \end{bmatrix}. \tag{11.1.3}$$

根据定理 3.2.2 得问题(11.1.1) 局部极小点的二阶充分条件:

若存在 $\bar{\boldsymbol{x}} \in \mathbb{R}^n$, $\bar{\boldsymbol{v}} \in \mathbb{R}^m$, 使得(11.1.3)式成立, 且对一切非零向量 $\boldsymbol{z} \in \mathbb{R}^n$, 只要 $\boldsymbol{A}\boldsymbol{z} = \boldsymbol{0}$, 便有 $\boldsymbol{z}^{\mathrm{T}}\boldsymbol{Q}\boldsymbol{z} > 0$, 则 $\bar{\boldsymbol{x}}$ 为问题(11.1.1)的严格局部极小点.

Lagrange 乘子法就是求解可行域内的 K－T 点, 即 Lagrange 函数的稳定点. 假设 $\boldsymbol{Q}$ 可逆, 考虑方程组

$$\begin{cases} Qx + c - A^{\mathrm{T}}v = 0, \\ Ax = b, \end{cases} \tag{11.1.4}$$

将第一个方程两边左乘 AQ^{-1}，再用第二个方程代入，得

$$(AQ^{-1}A^{\mathrm{T}})v = b - AQ^{-1}c,$$

求得此方程组的解 $\bar{v}$，再代回到第一个方程，求出问题(11.1.1)的 K－T 点

$$\bar{x} = Q^{-1}(A^{\mathrm{T}}\bar{v} - c).$$

如果 Q 为半正定矩阵，那么问题(11.1.1)的 K－T 点就是其全局极小点.

下面用分块矩阵的方法给出求问题(11.1.1)的 K－T 点的另一种形式.

若线性方程组(11.1.4)式的系数矩阵 $\begin{bmatrix} Q & -A^{\mathrm{T}} \\ -A & 0 \end{bmatrix}$ 可逆，则设

$$\begin{bmatrix} Q & -A^{\mathrm{T}} \\ -A & 0 \end{bmatrix}^{-1} = \begin{bmatrix} H & -R \\ -R^{\mathrm{T}} & G \end{bmatrix}. \tag{11.1.5}$$

从而由(11.1.4)式可得问题(11.1.1)的全局极小点

$$\begin{cases} \bar{x} = -Hc + Rb, \\ \bar{v} = R^{\mathrm{T}}c - Gb. \end{cases} \tag{11.1.6}$$

当 Q^{-1} 存在时，由(11.1.5)式可得 H、R、G 的表达式为

$$H = Q^{-1} - Q^{-1}A^{\mathrm{T}}(AQ^{-1}A^{\mathrm{T}})^{-1}AQ^{-1},$$

$$R = Q^{-1}A^{\mathrm{T}}(AQ^{-1}A^{\mathrm{T}})^{-1},$$

$$G = -(AQ^{-1}A^{\mathrm{T}})^{-1}.$$

设 x_0 是问题(11.1.1)的任一可行解，即 $Ax_0 = b$，在 x_0 处目标函数的梯度

$$\nabla f(x_0) = Qx_0 + c,$$

则(11.1.6)式可改写为

$$\begin{cases} \bar{x} = x_0 - H\nabla f(x_0), \\ \bar{v} = R^{\mathrm{T}}\nabla f(x_0). \end{cases} \tag{11.1.7}$$

例 11.1.1 用 Lagrange 乘子法求解问题

$$\begin{cases} \min & x_1^2 + 2x_2^2 + x_3^2 + x_3 - 2x_1x_2; \\ \text{s.t.} & x_1 + x_2 + x_3 = 4, \\ & 2x_1 - x_2 + x_3 = 2. \end{cases}$$

解 易知

$$Q = \begin{bmatrix} 2 & -2 & 0 \\ -2 & 4 & 0 \\ 0 & 0 & 2 \end{bmatrix}, c = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, A = \begin{bmatrix} 1 & 1 & 1 \\ 2 & -1 & 1 \end{bmatrix}, b = \begin{bmatrix} 4 \\ 2 \end{bmatrix}.$$

显然，Q 可逆且

$$Q^{-1}=\begin{bmatrix}1&\frac{1}{2}&0\\\frac{1}{2}&\frac{1}{2}&0\\0&0&\frac{1}{2}\end{bmatrix}.$$

因此,按上述 $\boldsymbol{H}$、$\boldsymbol{R}$、$\boldsymbol{G}$ 的表达式可得

$$\boldsymbol{H}=\frac{4}{11}\begin{bmatrix}\frac{1}{2}&\frac{1}{4}&-\frac{3}{4}\\\frac{1}{4}&\frac{1}{8}&-\frac{3}{8}\\-\frac{3}{4}&-\frac{3}{8}&\frac{9}{8}\end{bmatrix},\boldsymbol{R}=\frac{4}{11}\begin{bmatrix}\frac{3}{4}&\frac{3}{4}\\\frac{7}{4}&-1\\\frac{1}{4}&\frac{1}{4}\end{bmatrix},\boldsymbol{G}=-\frac{4}{11}\begin{bmatrix}3&-\frac{5}{2}\\-\frac{5}{2}&3\end{bmatrix},$$

从而,所求问题的全局极小点为

$$\bar{\boldsymbol{x}}=-\boldsymbol{Hc}+\boldsymbol{Rd}=\left(\frac{21}{11},\frac{43}{22},\frac{3}{22}\right)^{\mathrm{T}}.$$

□

11.1.2 直接消去法

求解问题(11.1.1)最简单而最直接的方法就是利用约束条件来消去部分变量,从而把问题转化成无约束问题,这一方法称为直接消去法(direct elimination method).

将矩阵 $\boldsymbol{A}$ 分解为 $\boldsymbol{A}=(\boldsymbol{B},\boldsymbol{N})$,其中 $\boldsymbol{B}\in\mathbb{R}^{m\times m}$为基矩阵,相应地,将 $\boldsymbol{x},\boldsymbol{c},\boldsymbol{Q}$ 作如下分块:

$$\boldsymbol{x}=\begin{bmatrix}\boldsymbol{x}_B\\\boldsymbol{x}_N\end{bmatrix},\boldsymbol{c}=\begin{bmatrix}\boldsymbol{c}_B\\\boldsymbol{c}_N\end{bmatrix},\boldsymbol{Q}=\begin{bmatrix}\boldsymbol{Q}_{11}&\boldsymbol{Q}_{12}\\\boldsymbol{Q}_{21}&\boldsymbol{Q}_{22}\end{bmatrix},$$

其中 $\boldsymbol{Q}_{11}\in\mathbb{R}^{m\times m}$. 这样,问题(11.1.1)的约束条件变成

$$\boldsymbol{Bx}_B+\boldsymbol{Nx}_N=\boldsymbol{b},$$

即得

$$\boldsymbol{x}_B=\boldsymbol{B}^{-1}\boldsymbol{b}-\boldsymbol{B}^{-1}\boldsymbol{Nx}_N,$$

代入$f(\boldsymbol{x})$中就得到与问题(11.1.1)等价的无约束问题

$$\min\ \varphi(\boldsymbol{x}_N)=\frac{1}{2}\boldsymbol{x}_N^{\mathrm{T}}\hat{\boldsymbol{Q}}_2\boldsymbol{x}_N+\hat{\boldsymbol{c}}_N^{\mathrm{T}}\boldsymbol{x}_N,\qquad(11.1.8)$$

其中

$$\hat{\boldsymbol{Q}}_2=\boldsymbol{Q}_{22}-\boldsymbol{Q}_{21}\boldsymbol{B}^{-1}\boldsymbol{N}-\boldsymbol{N}^{\mathrm{T}}(\boldsymbol{B}^{-1})^{\mathrm{T}}\boldsymbol{Q}_{12}+\boldsymbol{N}^{\mathrm{T}}(\boldsymbol{B}^{-1})^{\mathrm{T}}\boldsymbol{Q}_{11}\boldsymbol{B}^{-1}\boldsymbol{N},$$

$$\hat{\boldsymbol{c}}_N=\boldsymbol{c}_N-\boldsymbol{N}^{\mathrm{T}}(\boldsymbol{B}^{-1})^{\mathrm{T}}\boldsymbol{c}_B+(\boldsymbol{Q}_{21}-\boldsymbol{N}^{\mathrm{T}}(\boldsymbol{B}^{-1})^{\mathrm{T}}\boldsymbol{Q}_{11})\boldsymbol{B}^{-1}\boldsymbol{b}.$$

将函数 $\varphi(\boldsymbol{x}_N)$在点 $\bar{\boldsymbol{x}}_N\in\mathbb{R}^{n-m}$进行 Taylor 展开,可得

$$\varphi(\boldsymbol{x}_N)=\varphi(\bar{\boldsymbol{x}}_N)+\nabla\varphi(\bar{\boldsymbol{x}}_N)(\boldsymbol{x}_N-\bar{\boldsymbol{x}}_N)+\frac{1}{2}(\boldsymbol{x}_N-\bar{\boldsymbol{x}}_N)^{\mathrm{T}}\hat{\boldsymbol{Q}}_2(\boldsymbol{x}_N-\bar{\boldsymbol{x}}_N),$$

因此 $\bar{\boldsymbol{x}}_N$ 是 $\varphi(\boldsymbol{x}_N)$ 的全局极小点当且仅当矩阵 $\hat{\boldsymbol{Q}}_2$ 是半正定的，且

$$\nabla\varphi(\boldsymbol{x}_N)=\hat{\boldsymbol{Q}}_2\bar{\boldsymbol{x}}_N+\hat{\boldsymbol{c}}_N=\boldsymbol{0}.$$

如果 $\hat{\boldsymbol{Q}}_2$ 正定，那么问题(11.1.8)的全局极小点为

$$\bar{\boldsymbol{x}}_N=-\hat{\boldsymbol{Q}}_2^{-1}\hat{\boldsymbol{c}}_N. \tag{11.1.9}$$

从而问题(11.1.1)的全局极小点为

$$\bar{\boldsymbol{x}}=\begin{bmatrix}\bar{\boldsymbol{x}}_B\\ \bar{\boldsymbol{x}}_N\end{bmatrix}=\begin{bmatrix}\boldsymbol{B}^{-1}\boldsymbol{b}\\ \boldsymbol{0}\end{bmatrix}+\begin{bmatrix}\boldsymbol{B}^{-1}\boldsymbol{N}\\ -\boldsymbol{I}\end{bmatrix}\hat{\boldsymbol{Q}}_2^{-1}\hat{\boldsymbol{c}}_N.$$

记 $\bar{\boldsymbol{x}}$ 处的 Lagrange 乘子为 $\bar{\boldsymbol{v}}$，则有

$$\boldsymbol{A}^{\mathrm{T}}\bar{\boldsymbol{v}}=\nabla f(\bar{\boldsymbol{x}})=\boldsymbol{Q}\bar{\boldsymbol{x}}+\boldsymbol{c},$$

于是

$$\bar{\boldsymbol{v}}=(\boldsymbol{B}^{-1})^{\mathrm{T}}(\boldsymbol{Q}_{11}\bar{\boldsymbol{x}}_B+\boldsymbol{Q}_{12}\bar{\boldsymbol{x}}_N+\boldsymbol{c}_B).$$

如果 $\hat{\boldsymbol{Q}}_2$ 半正定且问题(11.1.8)有下界，那么问题(11.1.8)存在全局极小点；如果 $\hat{\boldsymbol{Q}}_2$ 半正定且问题(11.1.8)无下界，或 $\hat{\boldsymbol{Q}}_2$ 不是半正定的，那么问题(11.1.1)不存在有限解.

例 11.1.2 求解二次规划问题

$$\begin{cases}\min & f(\boldsymbol{x})=x_1^2+x_2^2+x_3^2;\\ \text{s.t.} & x_1+2x_2-x_3=4,\\ & x_1-x_2+x_3=-2.\end{cases}$$

解 将约束条件写成

$$\begin{cases}x_1+2x_2=4+x_3,\\ x_1-x_2=-2-x_3.\end{cases}$$

用 Gauss 消元法求得

$$x_1=-\frac{1}{3}x_3,\ x_2=2+\frac{2}{3}x_3.$$

并代入 $f(\boldsymbol{x})$ 中可得等价的无约束问题

$$\min\varphi(x_3)=\frac{14}{9}x_3^2+\frac{8}{3}x_3+4,$$

其中 $\hat{\boldsymbol{Q}}_2=\left(\frac{28}{9}\right)$，显然正定，故令 $\nabla\varphi(x_3)=0$，求得 $\bar{x}_3=-\frac{6}{7}$. 从而得 $\bar{x}_1=\frac{2}{7}$，$\bar{x}_2=\frac{10}{7}$. 因此，所求问题有惟一的全局极小点

$$\bar{\boldsymbol{x}}=(\bar{x}_1,\bar{x}_2,\bar{x}_3)^{\mathrm{T}}=\left(\frac{2}{7},\frac{10}{7},-\frac{6}{7}\right)^{\mathrm{T}}.$$

再利用 $A^{\mathrm{T}}\bar{v} = Q\bar{x} + c$,即

$$\begin{bmatrix} 1 & 1 \\ 2 & -1 \\ -1 & 1 \end{bmatrix}\begin{bmatrix} \bar{v}_1 \\ \bar{v}_2 \end{bmatrix} = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}\begin{bmatrix} 2/7 \\ 10/7 \\ -6/7 \end{bmatrix},$$

可求得 $\bar{v}_1 = \frac{8}{7}, \bar{v}_2 = -\frac{4}{7}$. □

直接消去法思想简单明了,使用方便. 不足之处是 B 可能接近一个奇异方阵,从而引起全局极小点 $\bar{x}$ 的数值不稳定.

11.1.3 广义消去法

广义消去法(general elimination method)并不需要假定 Q 可逆,因而能够处理一般的带等式约束的二次规划问题.

设 $A \in \mathbb{R}^{m\times n}$ 且 rank $A = m$, $Z \in \mathbb{R}^{n\times(n-m)}$ 的列向量组是齐次方程组 $Az = 0$ 的基础解系. 若存在矩阵 $Y \in \mathbb{R}^{n\times m}$,使得 (Y,Z) 可逆,则有 $AZ = 0$,且 $AY \in \mathbb{R}^{m\times m}$ 为非奇异矩阵. 于是问题(11.1.1)的约束方程 $Ax = b$ 的通解可表示为

$$x = Y(AY)^{-1}b + Z\eta, \qquad (11.1.10)$$

其中 η 为任意 $n-m$ 维实向量. 将(11.1.10)式代入问题(11.1.1),得到与问题(11.1.1)等价的无约束最优化问题

$$\min \psi(x) = \frac{1}{2}\eta^{\mathrm{T}}(Z^{\mathrm{T}}QZ)\eta + [c + QY(AY)^{-1}b]^{\mathrm{T}}Z\eta,$$

如果 $Z^{\mathrm{T}}QZ$ 正定,那么由(11.1.9)式可知,上述无约束最优化问题的全局极小点

$$\bar{\eta} = -(Z^{\mathrm{T}}QZ)^{-1}Z^{\mathrm{T}}[c + QY(AY)^{-1}b], \qquad (11.1.11)$$

将(11.1.11)式代入(11.1.10)式,得

$$\bar{x} = Y(AY)^{-1}b - Z(Z^{\mathrm{T}}QZ)^{-1}Z^{\mathrm{T}}[c + QY(AY)^{-1}b], \qquad (11.1.12)$$

再将(11.1.2)的第一式两边左乘 Y^{T},然后用(11.1.12)代入,则相应的 Lagrange 乘子可表示为

$$\begin{aligned} \bar{v} &= [(AY)^{-1}]^{\mathrm{T}}Y^{\mathrm{T}}(Q\bar{x} + c) \\ &= [(AY)^{-1}]^{\mathrm{T}}Y^{\mathrm{T}}[I - QZ(Z^{\mathrm{T}}QZ)^{-1}Z^{\mathrm{T}}]c + \\ &\quad [(AY)^{-1}]^{\mathrm{T}}Y^{\mathrm{T}}Q[I - Z(Z^{\mathrm{T}}QZ)^{-1}Z^{\mathrm{T}}Q]Y(AY)^{-1}b. \end{aligned} \qquad (11.1.13)$$

如果适当选取 Y,使得 $AY = I$,则(11.1.12)式和(11.1.13)式变成

$$\bar{x} = Yb - Z(Z^{\mathrm{T}}QZ)^{-1}Z^{\mathrm{T}}(c + QYb),$$

$$\bar{v} = Y^{\mathrm{T}}(Pc + QP^{\mathrm{T}}Yb),$$

其中 $P = I - QZ(Z^{\mathrm{T}}QZ)^{-1}Z^{\mathrm{T}}$.

对于矩阵 Y 和 Z 的不同选取方式,可得到问题(11.1.1)的不同计算方法.

第一种方法是利用矩阵的 QR 分解来选取矩阵 Y 和 Z. 假设

$$A^{\mathrm{T}} = Q\begin{bmatrix} R \\ 0 \end{bmatrix},$$

其中 Q 为正交矩阵，R 为非奇异的上三角矩阵，记 $Q = (Q_1, Q_2)$ $(Q_1 \in \mathbb{R}^{n \times m})$，且取 $Y = Q_1(R^{-1})^{\mathrm{T}}, Z = Q_2$，可以验证 $A = R^{\mathrm{T}}Q_1^{\mathrm{T}}, AY = I, AZ = 0$.

第二种选取方法是取

$$Y = \begin{bmatrix} B^{-1} \\ 0 \end{bmatrix}, Z = \begin{bmatrix} -B^{-1}N \\ I \end{bmatrix},$$

则利用广义消去法得到 $\bar{x}$ 和 $\bar{v}$ 的计算公式与直接消去法所得公式完全相同.

其实，可以任意选取 $V \in \mathbb{R}^{(n-m) \times n}$，使得 $\begin{bmatrix} A \\ V \end{bmatrix}$ 可逆，令

$$(Y, Z) = \begin{bmatrix} A \\ V \end{bmatrix}^{-1}, \tag{11.1.14}$$

则所选矩阵 Y 和 Z 必满足 $AY = I, AZ = 0$. V 的选取方式不同，可得到不同的 Y 和 Z，从而导致不同的计算 $\bar{x}$ 和 $\bar{v}$ 的计算公式. 正由于这个一般形式(11.1.14)，许多早期求解二次规划的方法都可以看作广义消去法的特殊情形.

11.2 起作用集方法

对于一般约束的二次规划问题的求解，很自然的想法是利用等式约束二次规划问题的求解方法来解决，Fletcher (1971) 提出起作用集方法 (active set method) 就是基于这种思想. 起作用集方法的关键是，在每次迭代中，都以已知的可行点为起点，将在该点起作用的约束作为等式约束，而对在该点不起作用的约束暂时不予考虑，这样，就将一般约束的二次规划问题转化为有限个仅带等式约束的二次规划问题.

考虑一般约束的二次规划问题

$$\begin{cases} \min & f(x) = \dfrac{1}{2}x^{\mathrm{T}}Qx + c^{\mathrm{T}}x; \\ \text{s.t.} & Ax \geqslant b, \\ & Ex = e. \end{cases} \tag{11.2.1}$$

其中 Q 为 n 阶半正定矩阵，$c \in \mathbb{R}^n, A \in \mathbb{R}^{m \times n}, E \in \mathbb{R}^{l \times n}, b \in \mathbb{R}^m, e \in \mathbb{R}^l$，且不妨设 $\operatorname{rank} A = m$, $\operatorname{rank} E = l$.

记问题(11.2.1)的约束条件为

$$\alpha_i^{\mathrm{T}} x \geqslant b_i (i \in I = \{1, 2, \cdots, m\}),$$

$$\beta_j^{\mathrm{T}} x = e_j (j \in J = \{1, 2, \cdots, l\}),$$

其中 α_i^{T} 是矩阵 A 的第 i 个行向量，β_j^{T} 是矩阵 E 的第 j 个行向量，$b = (b_1, b_2, \cdots,$

$b_m)^{\mathrm{T}}$, $\boldsymbol{e}=(e_1,e_2,\cdots,e_l)^{\mathrm{T}}$.

设 $\bar{\boldsymbol{x}}$ 是问题(11.2.1)的可行点,如果起作用约束指标集 $I(\bar{\boldsymbol{x}})\cup J$ 是已知的,则可以运用求解等式约束二次规划问题的方法,求得问题

$$\begin{cases}\min & f(\boldsymbol{x})=\dfrac{1}{2}\boldsymbol{x}^{\mathrm{T}}\boldsymbol{Q}\boldsymbol{x}+\boldsymbol{c}^{\mathrm{T}}\boldsymbol{x};\\ \text{s. t.} & \boldsymbol{\alpha}_i^{\mathrm{T}}\boldsymbol{x}=b_i(i\in I(\bar{\boldsymbol{x}})),\\ & \boldsymbol{\beta}_j^{\mathrm{T}}\boldsymbol{x}=e_j(j\in J),\end{cases}\tag{11.2.2}$$

的全局极小点,但是起作用集 $I(\bar{\boldsymbol{x}})\cup J$ 通常是未知的,如何确定起作用约束指标集是求解一般约束二次规划问题的关键.

下面首先讨论问题(11.2.1)与问题(11.2.2)的解的关系.

由于在 $\bar{\boldsymbol{x}}$ 附近,问题(11.2.1)的可行点必是问题(11.2.2)的可行点,因此当 $\bar{\boldsymbol{x}}$ 是问题(11.2.1)的局部极小点时,$\bar{\boldsymbol{x}}$ 也是问题(11.2.2)的局部极小点.

假设 $\bar{\boldsymbol{x}}$ 是问题(11.2.2)的局部极小点,且是问题(11.2.1)的可行点,当 Lagrange 乘子 $\bar{u}_i\geqslant 0\ (i\in I(\bar{\boldsymbol{x}}))$ 时,令 $\bar{u}_i=0(i\in I\backslash I(\bar{\boldsymbol{x}}))$,则 $\bar{\boldsymbol{x}}$ 是问题(11.2.1)的 K-T 点. 因为问题(11.2.1)是一个凸二次规划,所以其 K-T 点就是它的全局极小点.

设在求解问题(11.2.2)的第 k 次迭代中,将前一次迭代得到的可行点 $\boldsymbol{x}_k$ 作为起点,在该点起作用约束的指标集用 $S_k\cup J$ 表示(其中 $S_k\subseteq I(\boldsymbol{x}_k)$),且通常要求所选约束指标集对应的约束函数的梯度是线性无关的. 令 $\boldsymbol{d}=\boldsymbol{x}-\boldsymbol{x}_k$,则

$$\begin{aligned}f(\boldsymbol{x})&=\frac{1}{2}(\boldsymbol{d}+\boldsymbol{x}_k)^{\mathrm{T}}\boldsymbol{Q}(\boldsymbol{d}+\boldsymbol{x}_k)+\boldsymbol{c}^{\mathrm{T}}(\boldsymbol{d}+\boldsymbol{x}_k)\\&=\frac{1}{2}\boldsymbol{d}^{\mathrm{T}}\boldsymbol{Q}\boldsymbol{d}+\boldsymbol{d}^{\mathrm{T}}\boldsymbol{Q}\boldsymbol{x}_k+\frac{1}{2}\boldsymbol{x}_k^{\mathrm{T}}\boldsymbol{Q}\boldsymbol{x}_k+\boldsymbol{c}^{\mathrm{T}}\boldsymbol{d}+\boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}_k\\&=\frac{1}{2}\boldsymbol{d}^{\mathrm{T}}\boldsymbol{Q}\boldsymbol{d}+(\boldsymbol{Q}\boldsymbol{x}_k+\boldsymbol{c})^{\mathrm{T}}\boldsymbol{d}+\left(\frac{1}{2}\boldsymbol{x}_k^{\mathrm{T}}\boldsymbol{Q}\boldsymbol{x}_k+\boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}_k\right).\end{aligned}$$

于是问题(11.2.2)转化成求关于 $\boldsymbol{d}$ 的等价最优化问题:

$$\begin{cases}\min & \varphi(\boldsymbol{d})=\dfrac{1}{2}\boldsymbol{d}^{\mathrm{T}}\boldsymbol{Q}\boldsymbol{d}+(\boldsymbol{Q}\boldsymbol{x}_k+\boldsymbol{c})^{\mathrm{T}}\boldsymbol{d};\\ \text{s. t.} & \boldsymbol{\alpha}_i^{\mathrm{T}}\boldsymbol{d}=0,i\in S_k,\\ & \boldsymbol{\beta}_j^{\mathrm{T}}\boldsymbol{d}=0,j\in J.\end{cases}\tag{11.2.3}$$

利用解等式约束的二次规划问题的方法求解问题(11.2.3),其全局极小点记为 $\boldsymbol{d}_k$,相应的 Lagrange 乘子为 $u_i^{(k)}(i\in S_k)$,$v_j^{(k)}(j\in J)$. 根据不同情形,采取下列相应的步骤.

(1) $\boldsymbol{d}_k=\boldsymbol{0}$. 此时 $\boldsymbol{x}_k$ 是问题

$$\begin{cases} \min & f(\boldsymbol{x}) = \frac{1}{2}\boldsymbol{x}^{\mathrm{T}}\boldsymbol{Q}\boldsymbol{x} + \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}; \\ \text{s.t.} & \boldsymbol{\alpha}_i^{\mathrm{T}}\boldsymbol{x} = b_i(i \in S_k), \\ & \boldsymbol{\beta}_j^{\mathrm{T}}\boldsymbol{x} = e_j(j \in J), \end{cases} \tag{11.2.4}$$

的 K－T 点. 如果 $u_i^{(k)} \geqslant 0(\forall i \in S_k)$ 均成立，则 $\boldsymbol{x}_k$ 是问题(11.2.1)的 K－T 点. 否则，令 $i_k \in S_k$，使得

$$u_{i_k}^{(k)} = \min_{i \in S_k} u_i^{(k)} < 0, \tag{11.2.5}$$

取 $S_k = S_k \backslash \{i_k\}$，重新求解问题(11.2.3).

(2) $\boldsymbol{d}_k \neq \boldsymbol{0}$.

当 $\boldsymbol{x}_k + \boldsymbol{d}_k$ 是问题(11.2.1)的可行点时，在第 $k+1$ 次迭代中，取迭代点为 $\boldsymbol{x}_{k+1} = \boldsymbol{x}_k + \boldsymbol{d}_k$.

当 $\boldsymbol{x}_k + \boldsymbol{d}_k$ 不是可行点时，取搜索方向为 $\boldsymbol{d}_k$，并取迭代点

$$\boldsymbol{x}_{k+1} = \boldsymbol{x}_k + \lambda_k \boldsymbol{d}_k,$$

其中 λ_k 为步长. 按保持可行性的要求，步长 λ_k 的取值应使得对于任意 $i \notin S_k$ 都有

$$\boldsymbol{\alpha}_i^{\mathrm{T}}(\boldsymbol{x}_k + \lambda_k \boldsymbol{d}_k) \geqslant b_i. \tag{11.2.6}$$

由于 $\boldsymbol{x}_k$ 是可行点，$\boldsymbol{\alpha}_i^{\mathrm{T}}\boldsymbol{x}_k \geqslant b_i$，因此当 $\boldsymbol{\alpha}_i^{\mathrm{T}}\boldsymbol{d}_k \geqslant 0$ 时，对任意的非负数 λ_k，(11.2.6)式总成立；当 $\boldsymbol{\alpha}_i^{\mathrm{T}}\boldsymbol{d}_k < 0$ 时，只需取正数

$$\lambda_k = \min\left\{1, \min_{\substack{i \notin S_k \\ \boldsymbol{\alpha}_i^{\mathrm{T}}\boldsymbol{p}_k < 0}} \frac{b_i - \boldsymbol{\alpha}_i^{\mathrm{T}}\boldsymbol{d}_k}{\boldsymbol{\alpha}_i^{\mathrm{T}}\boldsymbol{d}_k}\right\} \tag{11.2.7}$$

即可. 如果 $\lambda_k = 1$，即

$$\frac{b_i - \boldsymbol{\alpha}_i^{\mathrm{T}}\boldsymbol{d}_k}{\boldsymbol{\alpha}_i^{\mathrm{T}}\boldsymbol{d}_k} \geqslant 1(\forall i \notin S_k),$$

则 $\boldsymbol{x}_k + \boldsymbol{d}_k$ 是问题(11.2.1)的可行点；

如果 $\lambda_k < 1$，即存在 $i_0 \notin S_k$，使得

$$\frac{b_{i_0} - \boldsymbol{\alpha}_{i_0}^{\mathrm{T}}\boldsymbol{d}_k}{\boldsymbol{\alpha}_{i_0}^{\mathrm{T}}\boldsymbol{d}_k} < 1,$$

则在点 $\boldsymbol{x}_{k+1}$ 处有

$$\boldsymbol{\alpha}_{i_0}^{\mathrm{T}}\boldsymbol{x}_{k+1} = \boldsymbol{\alpha}_{i_0}^{\mathrm{T}}(\boldsymbol{x}_k + \lambda_k \boldsymbol{d}_k) = b_{i_0},$$

这说明，在点 $\boldsymbol{x}_{k+1}$ 处，$\boldsymbol{\alpha}_{i_0}^{\mathrm{T}}\boldsymbol{x} \geqslant b_{i_0}$ 为起作用约束. 因此，应把指标 i_0 加入点 $\boldsymbol{x}_{k+1}$ 处的起作用约束指标集 S_{k+1} 中，即令 $S_{k+1} = S_k \cup \{i_0\}$.

下面给出起作用集方法的算法步骤.

算法 11－1(起作用集法)

Step 1　选取初始数据. 给定初始可行点 $\boldsymbol{x}_1$，相应的起作用约束指标集为

$S_1 \cup J$，令 $k=1$.

Step 2　求解等式约束问题. 求解问题(11.2.3)，设其全局极小点为 $\boldsymbol{d}_k$. 若 $\boldsymbol{d}_k=\boldsymbol{0}$，则转 Step 4；否则，由(11.2.7)式计算步长 λ_k，令 $\boldsymbol{x}_{k+1}:=\boldsymbol{x}_k+\lambda_k\boldsymbol{d}_k$，计算 $\nabla f(\boldsymbol{x}_{k+1})$，转 Step 3.

Step 3　修改起作用约束指标集. 若 $\lambda_k<1$，则存在 $i_0\notin S_k$，使 $\boldsymbol{\alpha}_{i_0}^{\mathrm{T}}\boldsymbol{x}_{k+1}=b_{i_0}$，故令 $S_k:=S_k\cup\{i_0\}$，置 $k:=k+1$，返回 Step 2；若 $\lambda_k=1$，则置 $k:=k+1$，转Step 4.

Step 4　检查是否满足终止准则. 用公式(11.1.7)计算出所有起作用约束对应的 Lagrange 乘子 $u_i^{(k)}(i\in S_k)$，$v_j^{(k)}(j\in J)$，令

$$u_{i_k}^{(k)}=\min_{i\in S_k}u_i^{(k)},$$

若 $u_{i_k}^{(k)}\geqslant 0$，则迭代终止，$\boldsymbol{x}_k$ 为问题(11.2.1)的全局极小点；否则，令 $S_k:=S_k\setminus\{i_k\}$，转 Step 2.

下面分析起作用集算法在第 k 次迭代后产生的全局极小点满足的性质.

定理 11.2.1　设在第 k 次迭代中，起作用约束指标集 $S_k\cup J$ 对应的约束函数的梯度是线性无关的，问题(11.2.3)全局极小点 $\boldsymbol{d}_k=\boldsymbol{0}$，

(1) 若存在 $i_k\in S_k$，使得 $u_{i_k}^{(k)}<0$，且等式约束二次规划问题

$$\begin{cases}\min & \varphi(\boldsymbol{d})=\dfrac{1}{2}\boldsymbol{d}^{\mathrm{T}}\boldsymbol{Q}\boldsymbol{d}+(\boldsymbol{Q}\boldsymbol{x}_k+\boldsymbol{c})^{\mathrm{T}}\boldsymbol{d};\\ \text{s.t.} & \boldsymbol{\alpha}_i^{\mathrm{T}}\boldsymbol{d}=0, i\in S_k\setminus\{i_k\},\\ & \boldsymbol{\beta}_j^{\mathrm{T}}\boldsymbol{d}=0, j\in J.\end{cases}\tag{11.2.8}$$

的全局极小点为 $\tilde{\boldsymbol{d}}_k$，则 $\boldsymbol{\alpha}_{i_k}^{\mathrm{T}}\tilde{\boldsymbol{d}}_k\geqslant 0$；

(2) 设问题(11.2.8)的约束矩阵为 $\boldsymbol{A}_k$，以方程组 $\boldsymbol{A}_k\boldsymbol{z}=\boldsymbol{0}$ 的基础解系为列向量组的矩阵记为 $\boldsymbol{Z}_k$，若矩阵 $\boldsymbol{Z}_k^{\mathrm{T}}\boldsymbol{Q}\boldsymbol{Z}_k$ 正定，则 $\boldsymbol{\alpha}_{i_k}^{\mathrm{T}}\tilde{\boldsymbol{d}}_k>0$，且 $\tilde{\boldsymbol{d}}_k$ 是问题(11.2.8)的目标函数的下降方向.

证明　(1) 由于 $\boldsymbol{x}_k$ 为问题(11.2.4)的 K－T 点，因此存在 Lagrange 乘子 $u_i^{(k)}(i\in S_k)$，$v_j^{(k)}(j\in J)$，使得

$$\boldsymbol{Q}\boldsymbol{x}_k+\boldsymbol{c}=\sum_{i\in S_k}u_i^{(k)}\boldsymbol{\alpha}_i+\sum_{j\in J}v_j^{(k)}\boldsymbol{\beta}_j.\tag{11.2.9}$$

又由 $\tilde{\boldsymbol{d}}_k$ 是问题(11.2.8)的全局极小点，可知存在 Lagrange 乘子 $\tilde{u}_i^{(k)}(i\in S_k\setminus\{i_k\})$，$v_j^{(k)}(j\in J)$，使得

$$\boldsymbol{Q}\tilde{\boldsymbol{d}}_k+(\boldsymbol{Q}\boldsymbol{x}_k+\boldsymbol{c})=\sum_{i\in S_k\setminus\{i_k\}}\tilde{u}_i^{(k)}\boldsymbol{\alpha}_i+\sum_{j\in J}\tilde{v}_j^{(k)}\boldsymbol{\beta}_j.\tag{11.2.10}$$

将(11.2.9)式代入(11.2.10)，得

$$\boldsymbol{Q}\tilde{\boldsymbol{d}}_k=\sum_{i\in S_k\setminus\{i_k\}}(\tilde{u}_i^{(k)}-u_i^{(k)})\boldsymbol{\alpha}_i+\sum_{j\in J}(\tilde{v}_j^{(k)}-v_j^{(k)})\boldsymbol{\beta}_j-u_{i_k}^{(k)}\boldsymbol{\alpha}_{i_k},\tag{11.2.11}$$

将(11.2.11)两边左乘 $\tilde{\boldsymbol{d}}_k^{\mathrm{T}}$,并利用问题(11.2.8)的约束条件,得

$$\tilde{\boldsymbol{d}}_k^{\mathrm{T}}\boldsymbol{Q}\tilde{\boldsymbol{d}}_k = -u_{i_k}^{(k)}\tilde{\boldsymbol{d}}_k^{\mathrm{T}}\boldsymbol{\alpha}_{i_k} = -u_{i_k}^{(k)}\boldsymbol{\alpha}_{i_k}^{\mathrm{T}}\tilde{\boldsymbol{d}}_k, \tag{11.2.12}$$

因 $\boldsymbol{Q}$ 为 n 阶半正定矩阵,且 $u_{i_k}^{(k)}<0$,故 $\boldsymbol{\alpha}_{i_k}^{\mathrm{T}}\tilde{\boldsymbol{d}}_k \geqslant 0$.

(2) 问题(11.2.8)的全局极小点 $\tilde{\boldsymbol{d}}_k$ 是方程组 $\boldsymbol{A}_k\boldsymbol{z}=\boldsymbol{0}$ 的解,则存在向量 $\boldsymbol{\eta}$,使得 $\tilde{\boldsymbol{d}}_k=\boldsymbol{Z}_k\boldsymbol{\eta}$,而 $\boldsymbol{Z}_k^{\mathrm{T}}\boldsymbol{Q}\boldsymbol{Z}_k$ 正定,从而

$$\tilde{\boldsymbol{d}}_k^{\mathrm{T}}\boldsymbol{Q}\tilde{\boldsymbol{d}}_k=\boldsymbol{\eta}^{\mathrm{T}}\boldsymbol{Z}_k^{\mathrm{T}}\boldsymbol{Q}\boldsymbol{Z}_k\boldsymbol{\eta}\geqslant 0,$$

且上式等于零的充要条件是 $\boldsymbol{\eta}=\boldsymbol{0}$,即 $\tilde{\boldsymbol{d}}_k=\boldsymbol{0}$. 又因约束指标集 $S_k\cup J$ 对应的约束函数的诸梯度是线性无关的,故由(11.2.11)式知,如果 $\tilde{\boldsymbol{d}}_k=\boldsymbol{0}$,那么 $u_{i_k}^{(k)}=0$,与条件 $u_{i_k}^{(k)}<0$ 矛盾,于是 $\tilde{\boldsymbol{d}}_k^{\mathrm{T}}\boldsymbol{Q}\tilde{\boldsymbol{d}}_k>0$,再由(11.2.12)有,$\boldsymbol{\alpha}_{i_k}^{\mathrm{T}}\tilde{\boldsymbol{d}}_k>0$.

将(11.2.10)两边左乘 $\tilde{\boldsymbol{d}}_k^{\mathrm{T}}$,得问题(11.2.8)的目标函数值

$$\varphi(\tilde{\boldsymbol{d}}_k)=-\frac{1}{2}\tilde{\boldsymbol{d}}_k^{\mathrm{T}}\boldsymbol{Q}\tilde{\boldsymbol{d}}_k<0,$$

说明 $\tilde{\boldsymbol{d}}_k$ 是问题(11.2.8)的下降方向. □

定理 11.2.2 假设在第 k 次迭代中,问题(11.2.3)的全局极小点 $\boldsymbol{d}_k\neq\boldsymbol{0}$,且满足定理 11.2.1(2)的条件,则 $\boldsymbol{d}_k$ 是问题(11.2.3)的目标函数的下降方向.

证明 因为矩阵 $\boldsymbol{Z}_k^{\mathrm{T}}\boldsymbol{Q}\boldsymbol{Z}_k$ 正定,所以 $\boldsymbol{d}_k\neq\boldsymbol{0}$ 为问题(11.2.3)的全局极小点. 同时 $\boldsymbol{d}=\boldsymbol{0}$ 是问题(11.2.3)的一个可行点,从而 $\varphi(\boldsymbol{d}_k)<\varphi(\boldsymbol{0})$,即

$$\frac{1}{2}\boldsymbol{d}_k^{\mathrm{T}}\boldsymbol{Q}\boldsymbol{d}_k+(\boldsymbol{Q}\boldsymbol{x}_k+\boldsymbol{c})^{\mathrm{T}}\boldsymbol{d}_k<0,$$

又由 $\boldsymbol{d}_k^{\mathrm{T}}\boldsymbol{Q}\boldsymbol{d}_k\geqslant 0$,知

$$\nabla f(\boldsymbol{x}_k)^{\mathrm{T}}\boldsymbol{d}_k=(\boldsymbol{Q}\boldsymbol{x}_k+\boldsymbol{c})^{\mathrm{T}}\boldsymbol{d}_k<0,$$

因此 $\boldsymbol{d}_k$ 是问题(11.2.3)的目标函数的下降方向. □

根据定理 11.2.1 和定理 11.2.2,算法 11-1 每次迭代后,问题(11.2.3)的全局最优解都是问题(11.2.1)的目标函数的下降方向,无论是增加一个约束还是删除一个约束,选取的起作用约束指标集 S_k 都不可能重复出现. 而约束条件的个数是有限的,故不同的 S_k 也有限,因此算法 11-1 必在有限步内求得问题(11.2.1)的全局极小点.

例 11.2.1 用起作用集法求解下列二次凸规划问题

$$\begin{cases}\min & f(\boldsymbol{x})=x_1^2-x_1x_2+2x_2^2-x_1-10x_2;\\ \text{s.t.} & -3x_1-2x_2\geqslant -6,\\ & x_1\geqslant 0,\\ & x_2\geqslant 0.\end{cases} \tag{11.2.13}$$

解 易知

$$\boldsymbol{Q}=\begin{bmatrix}2 & -1\\ -1 & 4\end{bmatrix},\quad \boldsymbol{c}=\begin{bmatrix}-1\\ -10\end{bmatrix}.$$

取初始可行点 $\boldsymbol{x}_1=(0,0)^{\mathrm{T}}$，相应的起作用约束指标集 $S_1=\{2,3\}$. 求解关于 $\boldsymbol{d}=(d_1,d_2)^{\mathrm{T}}$ 的等式约束二次规划问题

$$\begin{cases}\min & (d_1)^2-d_1d_2+2d_2^2-d_1-10d_2;\\ \text{s. t.} & d_1=0,\\ & d_2=0.\end{cases}$$

得全局极小点 $\boldsymbol{d}_1=(d_1^{(1)},d_2^{(1)})^{\mathrm{T}}=(0,0)^{\mathrm{T}}$. 因此，$\boldsymbol{x}_1$ 是对应的等式约束问题的全局极小点. 为判断 $\boldsymbol{x}_1$ 是否为问题(11.2.13)的全局极小点，需要计算 Lagrange 乘子.

由 $S_1=\{2,3\}$ 知

$$\boldsymbol{A}=\begin{bmatrix}1&0\\0&1\end{bmatrix},\quad \boldsymbol{b}=\begin{bmatrix}0\\0\end{bmatrix}.$$

按公式(11.1.7)，求得乘子 $u_2^{(1)}=-1, u_3^{(1)}=-10$，故 $\boldsymbol{x}_1$ 不是问题(11.2.13)的全局极小点.

将 $u_3^{(1)}$ 对应的问题(11.2.13)中的第 3 个约束的指标从 S_1 中去掉，即 $S_1=\{2\}$，再解问题

$$\begin{cases}\min & d_1^2-d_1d_2+2d_2^2-d_1-10d_2;\\ \text{s. t.} & d_1=0.\end{cases}$$

得全局极小点 $\boldsymbol{d}_1=(d_1^{(1)},d_2^{(1)})^{\mathrm{T}}=\left(0,\dfrac{5}{2}\right)^{\mathrm{T}}$.

计算迭代步长

$$\lambda_1=\min\left\{1,\frac{6}{5}\right\}=1.$$

令

$$\boldsymbol{x}_2=\boldsymbol{x}_1+\lambda_1\boldsymbol{d}_1=\begin{bmatrix}0\\0\end{bmatrix}+\begin{bmatrix}0\\5/2\end{bmatrix}=\begin{bmatrix}0\\5/2\end{bmatrix},$$

并求出

$$\nabla f(\boldsymbol{x}_2)=\left(-\frac{7}{2},0\right)^{\mathrm{T}}.$$

由于 $\lambda_1=1$，因此 $S_2=\{2\}$，并计算相应的 Lagrange 乘子. 此时 $\boldsymbol{A}=(1,0)$，$\boldsymbol{b}=\boldsymbol{0}$，故按公式(11.1.7)可得 $u_2^{(2)}=-\dfrac{7}{2}$，因此 $\boldsymbol{x}_2$ 也不是问题(11.2.13)的全局极小点.

从 S_2 中删除指标 2，则 $S_2=\varnothing$，解问题

$$\min d_1^2-d_1d_2+2d_2^2-\frac{7}{2}d_1;$$

得全局极小点 $\boldsymbol{d}_2=(d_1^{(2)},d_2^{(2)})^{\mathrm{T}}=\left(2,\dfrac{1}{2}\right)^{\mathrm{T}}$. 此时 $\lambda_2=\min\left\{1,\dfrac{1}{7}\right\}=\dfrac{1}{7}$,再令

$$\boldsymbol{x}_3=\boldsymbol{x}_2+\lambda_2\boldsymbol{d}_2=\left(\frac{2}{7},\frac{18}{7}\right)^{\mathrm{T}},$$

并求出

$$\nabla f(\boldsymbol{x}_3)=(-3,0)^{\mathrm{T}}.$$

显然,$S_3=\{1\}$,再解问题

$$\begin{cases}\min & d_1^2-d_1d_2+2d_2^2-3d_1;\\ \text{s.t.} & -3d_1-2d_2=0.\end{cases}$$

得全局极小点 $\boldsymbol{d}_3=(d_1^{(3)},d_2^{(3)})^{\mathrm{T}}=\left(\dfrac{3}{14},-\dfrac{9}{28}\right)^{\mathrm{T}}$. 且 $\lambda_3=\min\{1,8\}=1$. 令

$$\boldsymbol{x}_4=\boldsymbol{x}_3+\lambda_3\boldsymbol{d}_3=\left(\frac{1}{2},\frac{9}{4}\right)^{\mathrm{T}},$$

并求出

$$\nabla f(\boldsymbol{x}_4)=\left(-\frac{9}{4},-\frac{3}{2}\right)^{\mathrm{T}}.$$

在点 $\boldsymbol{x}_4$ 处,$S_4=\{1\}$,用公式(11.1.7)可算出 $u_1^{(4)}=\dfrac{3}{4}>0$. 因此,$\boldsymbol{x}_4=\left(\dfrac{1}{2},\dfrac{9}{4}\right)^{\mathrm{T}}$,即为所求的问题(11.2.13)的全局极小点. □

11.3 Wolfe 算法

单纯形法对于求解线性规划问题是一个很有用的方法. 本节将要介绍的 Wolfe 算法就是试图将线性规划中的单纯形法推广到二次凸规划问题.

考虑二次规划问题

$$\begin{cases}\min & f(\boldsymbol{x})=\dfrac{1}{2}\boldsymbol{x}^{\mathrm{T}}\boldsymbol{Q}\boldsymbol{x}+\boldsymbol{c}^{\mathrm{T}}\boldsymbol{x};\\ \text{s.t.} & \boldsymbol{A}\boldsymbol{x}=\boldsymbol{b},\\ & \boldsymbol{x}\geqslant\boldsymbol{0}.\end{cases}\tag{11.3.1}$$

其中 $\boldsymbol{Q}$ 为 n 阶半正定矩阵,$\boldsymbol{c}\in\mathbb{R}^n$,$\boldsymbol{b}\in\mathbb{R}^m$,$\boldsymbol{A}\in\mathbb{R}^{m\times n}$,且不妨设 $\operatorname{rank}\boldsymbol{A}=m$.

首先讨论问题(11.3.1)的最优性条件. 问题(11.3.1)对应的 Lagrange 函数为

$$L(\boldsymbol{x},\boldsymbol{v},\boldsymbol{u})=f(\boldsymbol{x})-\boldsymbol{v}^{\mathrm{T}}(\boldsymbol{A}\boldsymbol{x}-\boldsymbol{b})-\boldsymbol{u}^{\mathrm{T}}\boldsymbol{x}.$$

问题(11.3.1)的全局极小点即 K-T 点 $\bar{\boldsymbol{x}}$ 所满足的条件是,存在 $\bar{\boldsymbol{v}}\in\mathbb{R}^n$,$\bar{\boldsymbol{u}}\in\mathbb{R}^m$ 使得

$$\begin{cases} Q\bar{x} - A^{T}\bar{v} - \bar{u} = -c, \\ A\bar{x} = b, \\ \bar{u}^{T}\bar{x} = 0, \\ \bar{u} \geqslant 0, \bar{x} \geqslant 0 \end{cases} \tag{11.3.2}$$

成立. 若把问题(11.3.2)中的非线性部分 $u^{T}x$ 作为目标函数,式(11.3.2)中的线性部分作为约束条件,则可得到具有线性约束的非线性最优化问题:

$$\begin{cases} \min & u^{T}x; \\ \text{s.t.} & Qx - A^{T}v - u = -c, \\ & Ax = b, \\ & u \geqslant 0, x \geqslant 0. \end{cases} \tag{11.3.3}$$

设 $\begin{bmatrix} \bar{x} \\ \bar{v} \\ \bar{u} \end{bmatrix}$ 是问题(11.3.3)的全局极小点,若满足 $\bar{u}^{T}\bar{x} = 0$,则其中的 $\bar{x}$ 必为问题(11.3.1)的全局极小点. 因此,问题归结为求问题(11.3.3)的全局极小点.

为了使所有变量都非负,令 $v = v^{+} - v^{-}$,则问题(11.3.3)中的前两个约束条件成为

$$\begin{bmatrix} A & 0 & 0 & 0 \\ Q & -A^{T} & A^{T} & -I_{n} \end{bmatrix} \begin{bmatrix} x \\ v^{+} \\ v^{-} \\ u \end{bmatrix} = \begin{bmatrix} b \\ -c \end{bmatrix}, \tag{11.3.4}$$

后两个约束条件成为

$$x \geqslant 0, v^{+} \geqslant 0, v^{-} \geqslant 0, u \geqslant 0. \tag{11.3.5}$$

为了求问题(11.3.3)的全局极小点,我们先求(11.3.4)满足式(11.3.5)的基本解,为此特引入人工变量 y. 具体步骤如下:

(1) 先用单纯形法的两阶段法求得满足

$$Ax = b, x \geqslant 0$$

的解. 将 A 分解为 $A = (B, N)$,其中 B 为可行基,所得解的基变量记为 x_{B},非基变量 $x_{N} = 0$.

(2) 确定对角阵 $E = (\Delta_{j}\delta_{ij})$,其中 Δ_{j} 为 1 或 -1,δ_{ij} 满足

$$\delta_{ij} = \begin{cases} 1, & i = j, \\ 0, & i \neq j. \end{cases}$$

其意义是在式(11.3.4)中引入人工变量 y 后得到

$$\begin{bmatrix} \boldsymbol{A} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{Q} & -\boldsymbol{A}^{\mathrm{T}} & \boldsymbol{A}^{\mathrm{T}} & -\boldsymbol{I}_n & \boldsymbol{E} \end{bmatrix} \begin{bmatrix} \boldsymbol{x} \\ \boldsymbol{v}^{+} \\ \boldsymbol{v}^{-} \\ \boldsymbol{u} \\ \boldsymbol{y} \end{bmatrix} = \begin{bmatrix} \boldsymbol{b} \\ -\boldsymbol{c} \end{bmatrix}, \tag{11.3.6}$$

且当 $\boldsymbol{x}$ 用(1)中求得的解代入时,若 $\boldsymbol{v}^{+}=\boldsymbol{v}^{-}=\boldsymbol{u}=\boldsymbol{0}$,则必有 $\boldsymbol{y}\geqslant\boldsymbol{0}$. 将 $\boldsymbol{Q}$ 分解为 $\boldsymbol{Q}=(\boldsymbol{Q}_B,\boldsymbol{Q}_N)$,则 $\boldsymbol{E}$ 应满足

$$\boldsymbol{Q}_B\boldsymbol{x}_B+\boldsymbol{E}\boldsymbol{y}=-\boldsymbol{c}, \tag{11.3.7}$$

且其中的人工变量 $\boldsymbol{y}=(y_1,y_2,\cdots,y_n)^{\mathrm{T}}$ 非负.

记 $\boldsymbol{Q}_B$ 中的列向量依次为 $\boldsymbol{q}_1,\boldsymbol{q}_2,\cdots,\boldsymbol{q}_m$,则式(11.3.7)等价于

$$\Delta_j y_j=-(c_j+\boldsymbol{q}_j^{\mathrm{T}}\boldsymbol{x}_B),j=1,2,\cdots,n.$$

因此,$\boldsymbol{E}$ 中 Δ_j 的选取规则为

$$\Delta_j=\begin{cases} -1, & c_j+\boldsymbol{q}_j^{\mathrm{T}}\boldsymbol{x}_B>0, \\ 1, & \text{否则}. \end{cases}$$

(3) 在约束(11.3.6)和约束

$$\boldsymbol{x}\geqslant\boldsymbol{0},\boldsymbol{v}^{+}\geqslant\boldsymbol{0},\boldsymbol{v}^{-}\geqslant\boldsymbol{0},\boldsymbol{u}\geqslant\boldsymbol{0},\boldsymbol{y}\geqslant\boldsymbol{0},\boldsymbol{u}^{\mathrm{T}}\boldsymbol{x}=0 \tag{11.3.8}$$

下极小化线性目标函数 $\sum_{i=1}^{n} y_i$.

对于上述线性规划问题,极小化可采用单纯形法. 但必须注意,进行旋转变换时要求 $\boldsymbol{u}^{\mathrm{T}}\boldsymbol{x}=0$,即当某个 $x_j>0$ 时,u_j 不允许进入基变量. 旋转变换直到 $\sum_{i=1}^{n} y_i=0$ 时结束,此时得到的 $\bar{\boldsymbol{x}}$ 即为二次凸规划问题(11.3.1)的全局极小点.

例 11.3.1 用 Wolfe 算法求解二次凸规划问题

$$\begin{cases} \min & f=-2x_1-x_2+x_1^2; \\ \text{s.t.} & 2x_1+3x_2+x_3=6, \\ & 2x_1+x_2+x_4=5, \\ & x_1,x_2,x_3,x_4\geqslant 0. \end{cases}$$

解 易知

$$\boldsymbol{Q}=\begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad \boldsymbol{c}=\begin{bmatrix} -2 \\ -1 \\ 0 \\ 0 \end{bmatrix},$$

$$\boldsymbol{A}=\begin{bmatrix} 2 & 3 & 1 & 0 \\ 2 & 1 & 0 & 1 \end{bmatrix}, \quad \boldsymbol{b}=\begin{bmatrix} 6 \\ 4 \end{bmatrix},$$

首先确定 $\boldsymbol{Ax}=\boldsymbol{b},\boldsymbol{x}\geqslant\boldsymbol{0}$ 的一个解，显然可取

$$\boldsymbol{x}=\left(\frac{3}{2},1,0,0\right)^{\mathrm{T}},$$

再确定 $\boldsymbol{E}=(\Delta_j\delta_{ij})$. 由于

$$\boldsymbol{c}+\boldsymbol{Q}_B\boldsymbol{x}_B=\begin{bmatrix}-2\\-1\\0\\0\end{bmatrix}+\begin{bmatrix}2&0\\0&0\\0&0\\0&0\end{bmatrix}\begin{bmatrix}\frac{3}{2}\\1\end{bmatrix}=\begin{bmatrix}1\\-1\\0\\0\end{bmatrix},$$

因此 $\Delta_1=-1,\Delta_2=\Delta_3=\Delta_4=1$，从而

$$\boldsymbol{E}=\begin{bmatrix}-1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&1\end{bmatrix}.$$

下面利用单纯形法在约束(11.3.6)和(11.3.8)下极小化目标函数 $z=\sum_{i=1}^{4}y_i$. 表 11.3.1 中给出了初始数据，为了得到初始单纯形表，表 11.3.1 的最后一行对应于目标函数 z 中变量系数的相反数.

表 11.3.1　例 11.3.1 初始数据表

	x_1	x_2	x_3	x_4	v_1^+	v_2^+	v_1^-	v_2^-	u_1	u_2	u_3	u_4	y_1	y_2	y_3	y_4	
x_1	2	3	1	0	0	0	0	0	0	0	0	0	0	0	0	0	6
x_2	2	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	4
y_1	2	0	0	0	−2	−2	2	2	−1	0	0	0	−1	0	0	0	2
y_2	0	0	0	0	−3	−1	3	1	0	−1	0	0	0	1	0	0	1
y_3	0	0	0	0	−1	0	1	0	0	0	−1	0	0	0	1	0	0
y_4	0	0	0	0	0	−1	0	1	0	0	0	−1	0	0	0	1	0
z	0	0	0	0	0	0	0	0	0	0	0	0	−1	−1	−1	−1	0

经过初等行变换得到以 x_1,x_2,y_1,y_2,y_3,y_4 为基变量的初始单纯形表 11.3.2.

表 11.3.2　例 11.3.1 第一轮迭代后的初始单纯形表

	x_1	x_2	x_3	x_4	v_1^+	v_2^+	v_1^-	v_2^-	u_1	u_2	u_3	u_4	y_1	y_2	y_3	y_4	
x_1	1	0	$-\frac{1}{4}$	$\frac{3}{4}$	0	0	0	0	0	0	0	0	0	0	0	0	$\frac{3}{2}$
x_2	0	1	$\frac{1}{2}$	$-\frac{1}{2}$	0	0	0	0	0	0	0	0	0	0	0	0	1
y_1	0	0	$-\frac{1}{2}$	$\frac{3}{2}$	2	2	-2	-2	1	0	0	0	1	0	0	0	1
y_2	0	0	0	0	-3	-1	3	1	0	-1	0	0	0	1	0	0	1
y_3	0	0	0	0	-1	0	1^*	0	0	0	-1	0	0	0	1	0	0
y_4	0	0	0	0	0	-1	0	1	0	0	0	-1	0	0	0	1	0
z	0	0	$-\frac{1}{2}$	$\frac{3}{2}$	-2	0	2	0	1	-1	-1	-1	0	0	0	0	-2

由于 v_1^- 对应的检验数 2 最大，因此以 v_1^- 为进基变量，y_3 为离基变量，进行{5,7}旋转变换，得到单纯形表 11.3.3.

表 11.3.3　例 11.3.1 第二轮迭代后的单纯形表

	x_1	x_2	x_3	x_4	v_1^+	v_2^+	v_1^-	v_2^-	u_1	u_2	u_3	u_4	y_1	y_2	y_3	y_4	
x_1	1	0	$-\frac{1}{4}$	$\frac{3}{4}$	0	0	0	0	0	0	0	0	0	0	0	0	$\frac{3}{2}$
x_2	0	1	$\frac{1}{2}$	$-\frac{1}{2}$	0	0	0	0	0	0	0	0	0	0	0	0	1
y_1	0	0	$-\frac{1}{2}$	$\frac{3}{2}^*$	0	2	0	-2	1	0	-2	0	1	0	2	0	1
y_2	0	0	0	0	0	-1	0	1	0	-1	3	0	0	1	-3	0	1
v_1^-	0	0	0	0	-1	0	1	0	0	0	-1	0	0	0	1	0	0
y_4	0	0	0	0	0	-1	0	1	0	0	0	-1	0	0	0	1	0
z	0	0	$-\frac{1}{2}$	$\frac{3}{2}$	0	0	0	0	1	-1	1	-1	0	0	-2	-1	-2

现在是 x_4 对应的检验数最大，故以 x_4 为进基变量，y_1 为离基变量，进行{3,4}旋转变换，得到单纯形表 11.3.4.

表 11.3.4 例 11.3.1 第三轮迭代后的单纯形表

	x_1	x_2	x_3	x_4	v_1^+	v_2^+	v_1^-	v_2^-	u_1	u_2	u_3	u_4	y_1	y_2	y_3	y_4	
x_1	1	0	0	0	0	-1	0	1	$-\frac{1}{2}$	0	1	0	$-\frac{1}{2}$	0	-1	0	1
x_2	0	1	$\frac{1}{3}$	0	0	$\frac{2}{3}$	0	$-\frac{2}{3}$	$\frac{1}{3}$	0	$-\frac{2}{3}$	0	$\frac{1}{3}$	0	$\frac{2}{3}$	0	$\frac{4}{3}$
x_4	0	0	$-\frac{1}{3}$	1	0	$\frac{4}{3}$	0	$-\frac{4}{3}$	$\frac{2}{3}$	0	$-\frac{4}{3}$	0	$\frac{2}{3}$	0	$\frac{4}{3}$	0	$\frac{2}{3}$
y_2	0	0	0	0	0	-1	0	1	0	-1	3^*	0	0	1	-3	0	1
v_1^-	0	0	0	0	-1	0	1	0	0	0	-1	0	0	0	1	0	0
y_4	0	0	0	0	0	-1	0	1	0	0	0	-1	0	0	0	1	0
z	0	0	0	0	0	-2	0	2	0	-1	3	-1	-1	0	-4	-1	-1

由于 u_3 对应的检验数最大,因此以 u_3 为进基变量,y_2 为离基变量,进行{4,11}旋转变换,得到单纯形表 11.3.5.

表 11.3.5 例 11.3.1 第四轮迭代后的单纯形表

	x_1	x_2	x_3	x_4	v_1^+	v_2^+	v_1^-	v_2^-	u_1	u_2	u_3	u_4	y_1	y_2	y_3	y_4	
x_1	1	0	0	0	0	$-\frac{2}{3}$	0	$\frac{2}{3}$	$-\frac{1}{2}$	$\frac{1}{3}$	0	0	$-\frac{1}{2}$	$-\frac{1}{3}$	0	0	$\frac{2}{3}$
x_2	0	1	$\frac{1}{3}$	0	0	$\frac{4}{9}$	0	$-\frac{4}{9}$	$\frac{1}{3}$	$-\frac{2}{9}$	0	0	$\frac{1}{3}$	$\frac{2}{9}$	0	0	$\frac{14}{9}$
x_4	0	0	$-\frac{1}{3}$	1	0	$\frac{8}{9}$	0	$-\frac{8}{9}$	$\frac{2}{3}$	$\frac{4}{9}$	0	0	$\frac{2}{3}$	$\frac{4}{9}$	0	0	$\frac{10}{9}$
u_3	0	0	0	0	0	$-\frac{1}{3}$	0	$\frac{1}{3}$	0	$-\frac{1}{3}$	1	0	0	$\frac{1}{3}$	-1	0	$\frac{1}{3}$
v_1^-	0	0	0	0	-1	$-\frac{1}{3}$	1	$\frac{1}{3}$	0	$-\frac{1}{3}$	0	0	0	$\frac{1}{3}$	0	0	$\frac{1}{3}$
y_4	0	0	0	0	0	-1	0	1	0	0	0	-1	0	0	0	1	0
z	0	0	0	0	0	-1	0	1	0	0	0	-1	-1	-1	-1	0	0

从表 11.3.5 右下角的数值知,此时目标函数

$$z = \sum_{i=1}^{4} y_i = 0,$$

因此,该二次凸规划问题的全局极小点为

$$\bar{x} = \left(\frac{2}{3},\frac{14}{9},0,\frac{10}{9}\right)^{\mathrm{T}}.$$ □

最后,我们给出下面的重要结论.

定理 11.3.1 若问题(11.3.1)中二次目标函数的 Hesse 矩阵 $\boldsymbol{Q}$ 是正定的,则用 Wolfe 算法必可求得其全局极小点,即在 $\boldsymbol{u}^{\mathrm{T}}\boldsymbol{x}=0$ 条件下用单纯形法求解问题

$$\min\left\{\sum_{i=1}^{n} y_i \mid (11.3.6)\text{ 式},(11.3.8)\text{ 式}\right\}$$

时,必能使 $\sum_{i=1}^{n} y_i = 0$. □

一般来说,当 $\boldsymbol{Q}$ 是半正定阵时,Wolfe 算法有失败的可能.这时,可将主对角线上的元素作微小扰动,以 $\boldsymbol{Q}+\varepsilon\boldsymbol{I}(\varepsilon>0$ 很小$)$取代 $\boldsymbol{Q}$ 后,再利用 Wolfe 算法.

11.4 Lemke 算法

上一节介绍的 Wolfe 算法是最早出现的利用单纯形法求解二次规划的方法,本节将要介绍的 Lemke 算法是利用单纯形法求解二次规划的另一种方法.其基本思想是:把线性规划中的单纯形法加以适当修改,再用来求二次规划的 K-T 点.

考虑二次规划问题

$$\begin{cases}\min & f(\boldsymbol{x}) = \dfrac{1}{2}\boldsymbol{x}^{\mathrm{T}}\boldsymbol{Q}\boldsymbol{x} + \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x};\\ \text{s. t.} & \boldsymbol{A}\boldsymbol{x} \geqslant \boldsymbol{b},\\ & \boldsymbol{x} \geqslant \boldsymbol{0}.\end{cases} \tag{11.4.1}$$

其中 $\boldsymbol{Q}\in\mathbb{R}^{n\times n}$ 且对称,$\boldsymbol{c}\in\mathbb{R}^{n}$,$\boldsymbol{b}\in\mathbb{R}^{m}$,$\boldsymbol{A}\in\mathbb{R}^{m\times n}$,不妨设 $\mathrm{rank}\boldsymbol{A}=m$.

引入乘子 $\boldsymbol{u}$ 和 $\boldsymbol{v}$,定义 Lagrange 函数

$$L(\boldsymbol{x},\boldsymbol{u},\boldsymbol{v}) = f(\boldsymbol{x}) - \boldsymbol{u}^{\mathrm{T}}(\boldsymbol{A}\boldsymbol{x}-\boldsymbol{b}) - \boldsymbol{v}^{\mathrm{T}}\boldsymbol{x},$$

引入松弛变量(向量)$\boldsymbol{y}\geqslant\boldsymbol{0}$,使

$$\boldsymbol{A}\boldsymbol{x} - \boldsymbol{y} = \boldsymbol{b},$$

由定理 3.3.4 知,问题(11.4.1)的 K-T 条件为

$$\begin{cases}\boldsymbol{Q}\boldsymbol{x} - \boldsymbol{A}^{\mathrm{T}}\boldsymbol{u} - \boldsymbol{v} = -\boldsymbol{c},\\ \boldsymbol{A}\boldsymbol{x} - \boldsymbol{y} = \boldsymbol{b},\\ \boldsymbol{v}^{\mathrm{T}}\boldsymbol{x} = 0,\\ \boldsymbol{u}^{\mathrm{T}}\boldsymbol{y} = 0,\\ \boldsymbol{u}\geqslant\boldsymbol{0},\boldsymbol{v}\geqslant\boldsymbol{0},\boldsymbol{x}\geqslant\boldsymbol{0}.\end{cases} \tag{11.4.2}$$

令

$$w=\begin{bmatrix}v\\y\end{bmatrix},\quad z=\begin{bmatrix}x\\u\end{bmatrix},\quad M=\begin{bmatrix}Q & -A^{\mathrm{T}}\\A & 0\end{bmatrix},\quad q=\begin{bmatrix}c\\-b\end{bmatrix},$$

这里 $w\in\mathbb{R}^{m+n}$，$z\in\mathbb{R}^{m+n}$，$q\in\mathbb{R}^{m+n}$和 $M\in\mathbb{R}^{(m+n)\times(m+n)}$. 则(11.4.2)式的矩阵表示形式为

$$\begin{cases}w-Mz=q,\\w\geqslant 0,z\geqslant 0,\end{cases}\tag{11.4.3}$$

及

$$w^{\mathrm{T}}z=0.\tag{11.4.4}$$

求满足(11.4.3)式和(11.4.4)式的解 $\begin{bmatrix}w\\z\end{bmatrix}$ 的问题称为线性互补问题(linear complementary problem)，简称为 LCP，(11.4.4)式称为互补条件(complementary condition). 显然(11.4.4) 式等价于

$$w_jz_j=0(j=1,2,\cdots,m+n).\tag{11.4.4a}$$

设 $\begin{bmatrix}w\\z\end{bmatrix}$ 是线性互补问题的一个解，则由式(11.4.4a)知，其 $2(m+n)$ 个分量中至少有 $m+n$ 个分量等于 0，而且每个互补变量对 (w_j,z_j) 中至少有一个变量取零值. 这就是下面定义的互补基本可行解(complementary basic feasible solution).

定义 11.4.1 设 $\begin{bmatrix}w\\z\end{bmatrix}$ 是问题(11.4.3)的一个基本可行解，且每个互补变量对 (w_j,z_j) $(j=1,\cdots,m+n)$ 中只有一个变量是基变量，则称 $\begin{bmatrix}w\\z\end{bmatrix}$ 是线性互补问题(11.4.3)和(11.4.4a)的一个互补基本可行解，简称 CBF 解.

这样，求二次规划 K－T 点的问题就转化为求线性互补问题的 CBF 解. 下面，介绍求 CBF 解的 Lemke 算法. 分两种情况讨论：

(1) 若 $q\geqslant 0$，则 $\begin{bmatrix}w\\z\end{bmatrix}=\begin{bmatrix}q\\0\end{bmatrix}$ 就是一个 CBF 解；

(2) $q\not\geqslant 0$，则引入人工变量 z_0，使式(11.4.3)变成

$$\begin{cases}w-Mz-z_0e=q,\\z_0\geqslant 0,w_j\geqslant 0,z_j\geqslant 0,j=1,\cdots,m+n.\end{cases}\tag{11.4.5}$$

其中 $e=(1,1,\cdots,1)^{\mathrm{T}}\in\mathbb{R}^{m+n}$. 显然，若 $\begin{bmatrix}w\\z\\z_0\end{bmatrix}$ 是线性互补问题(11.4.4a)和(11.4.5)的可行解且 $z_0=0$，则 $\begin{bmatrix}w\\z\end{bmatrix}$ 就是所求的 CBF 解.

为叙述方便起见，我们先引入准互补基本可行解(almost complementary basic

feasible solution)的概念.

定义 11.4.2 设$\begin{bmatrix} w \\ z \\ z_0 \end{bmatrix}$是线性互补问题(11.4.4a)和(11.4.5)的可行解且满足

(1) $\begin{bmatrix} w \\ z \\ z_0 \end{bmatrix}$是(11.4.5)的基本可行解;

(2) 存在 $s \in \{1, \cdots, m+n\}$,使得 w_s 和 z_s 都不是基变量;

(3) 每个互补变量对(w_j, z_j) $(j = 1, 2, \cdots, m+n, j \neq s)$中恰有一个变量是基变量且 z_0 是基变量,则称$\begin{bmatrix} w \\ z \\ z_0 \end{bmatrix}$为线性互补问题(11.4.4a)和(11.4.5)的准互补基本可行解,简称 ACBF 解.

下面,用单纯形法求 ACBF 解.

首先,若令

$$z_0 = \max\{-q_j \mid j = 1, 2, \cdots, m+n\} = -q_s > 0,$$

$$\boldsymbol{z} = \boldsymbol{0}, \quad \boldsymbol{w} = \boldsymbol{q} - q_s \boldsymbol{e} \geqslant \boldsymbol{0},$$

则$\begin{bmatrix} w \\ z \\ z_0 \end{bmatrix}$是一个 ACBF 解,其中 $w_j (j \neq s)$ 和 z_0 是基变量,其余变量为非基变量.

以上述$\begin{bmatrix} w \\ z \\ z_0 \end{bmatrix}$为初始解,用单纯形法求新的 ACBF 解,力图通过这种方法迫使 z_0 变成非基变量. 为保持可行性,选择主元时要求遵守下面两条规则:

(1) 保持可行性:按照单纯形法中的 Danzig 规则确定离基变量;

(2) 保持准互补性:若 w_j(或 z_j)离基,则 z_j(或 w_j)进基.

这样将实现从一个 ACBF 解到另一个 ACBF 解的转换,直至得到一个 CBF 解,即 z_0 变成非基变量,或者判定由约束(11.4.4a)和(11.4.5)所定义的可行域无界.

下面,给出求 CBF 解的 Lemke 算法的具体步骤.

算法 11-2(Lemke 算法)

Step 1 若 $\boldsymbol{q} \geqslant \boldsymbol{0}$,则停止计算,$\begin{bmatrix} w \\ z \end{bmatrix} = \begin{bmatrix} q \\ 0 \end{bmatrix}$是 CBF 解;否则,用类似于单纯形

表的表格形式表示出式(11.4.5)中的第一个约束. 记

$$-q_s=\max\{-q_j \mid j=1,\cdots,m+n\},$$

取第 s 行为主行,z_0 对应的列为主列,将 z_0 换入基变量取代 w_s,对初始表进行旋转变换后得新表并令 $y_s=z_s$,转 Step 2.

Step 2　设当前表中变量 y_s 即变量 z_s 下面的列为 $\boldsymbol{d}_s$,若 $\boldsymbol{d}_s\leqslant\boldsymbol{0}$,则停止计算,得到问题(11.4.5)的可行域的极方向;否则,按最小比值规则确定指标 r,使

$$\frac{p_r}{d_r^{(s)}}=\min\left\{\frac{p_j}{d_j^{(s)}}\ \middle|\ d_j^{(s)}>0,j=1,\cdots,m+n\right\},$$

其中 p_j 为当前表格中右端第 j 个元素,即第 j 个基变量的当前值,$d_j^{(s)}$ 为向量 $\boldsymbol{d}_s$ 的第 j 个分量. 若第 r 行对应的基变量为 z_0,则转 Step 4;否则,转 Step 3.

Step 3　这时第 r 行的基变量为 w_l 或 $z_l(l\neq s)$,将该基变量作为离基变量,y_s 作为进基变量,进行旋转变换后得新表. 若离基变量是 w_l,则令 $y_s=z_l$;若离基变量是 z_l,则令 $y_s=w_l$,转 Step 2.

Step 4　这时,变量 y_s 进基,变量 z_0 出基,进行旋转变换后,所得新表中 z_0 为非基变量,停止计算,得到一个 CBF 解.

例 11.4.1　用 Lemke 算法求解例 11.2.1:

$$\begin{cases}\min & f(\boldsymbol{x})=x_1^2-x_1x_2+2x_2^2-x_1-10x_2;\\ \text{s.t.} & -3x_1-2x_2\geqslant-6,\\ & x_1\geqslant0,x_2\geqslant0.\end{cases}\tag{11.4.6}$$

解　易知

$$\boldsymbol{Q}=\begin{bmatrix}2&-1\\-1&4\end{bmatrix},\boldsymbol{c}=\begin{bmatrix}-1\\-10\end{bmatrix},\boldsymbol{A}=(-3,-2),b=-6,$$

$$\boldsymbol{M}=\begin{bmatrix}\boldsymbol{Q}&-\boldsymbol{A}^{\mathrm{T}}\\\boldsymbol{A}&0\end{bmatrix}=\begin{bmatrix}2&-1&3\\-1&4&2\\-3&-2&0\end{bmatrix},\boldsymbol{q}=\begin{bmatrix}\boldsymbol{c}\\-\boldsymbol{b}\end{bmatrix}=\begin{bmatrix}-1\\-10\\6\end{bmatrix}.$$

相应的线性互补问题为

$$\begin{cases}w_1-2x_1+x_2-3u=-1,\\ w_2+x_1-4x_2-2u=-10,\\ w_3+3x_1+2x_2=6,\\ w_1,w_2,w_3,x_1,x_2,u\geqslant0,\\ w_1x_1=w_2x_2=w_3u=0.\end{cases}\tag{11.4.7}$$

引入人工变量 z_0,建立表 11.4.1.

表 11.4.1　例 11.4.1 的初始表

	w_1	w_2	w_3	x_1	x_2	u	z_0	
w_1	1	0	0	-2	1	-3	-1	-1
w_2	0	1	0	1	-4	-2	-1^*	-10
w_3	0	0	1	3	2	0	-1	6

因 $\max\{-q_j \mid 1\leqslant j\leqslant 3\}=10=-q_2$，故选 w_2 出基，而 z_0 进基，旋转变换后得表 11.4.2.

表 11.4.2　例 11.4.1 迭代一次的表

	w_1	w_2	w_3	x_1	x_2	u	z_0	
w_1	1	-1	0	3	5^*	-1	0	9
z_0	0	-1	0	-1	4	2	1	10
w_3	0	-1	1	2	6	2	0	16

令 $y_s=z_2=x_2$，由于 $\boldsymbol{d}_s=(5,4,6)^{\mathrm{T}}\nleqslant\mathbf{0}$，故求出

$$\frac{p_r}{d_r^{(s)}}=\min\left\{\frac{9}{5},\frac{10}{4},\frac{16}{6}\right\}=\frac{9}{5}=\frac{p_1}{d_1^{(s)}}$$

即 $r=1$，选第一行的基变量 w_1 出基，$y_s=x_2$ 进基，旋转变换后得表 11.4.3.

表 11.4.3　例 11.4.1 迭代二次的表

	w_1	w_2	w_3	x_1	x_2	u	z_0	
x_2	$\frac{1}{5}$	$-\frac{1}{5}$	0	$-\frac{3}{5}$	1	$-\frac{1}{5}$	0	$\frac{9}{5}$
z_0	$-\frac{4}{5}$	$-\frac{1}{5}$	0	$\frac{7}{5}$	0	$\frac{14}{5}$	1	$\frac{14}{5}$
w_3	$-\frac{6}{5}$	$\frac{1}{5}$	1	$\frac{28}{5}^*$	0	$\frac{16}{5}$	0	$\frac{26}{5}$

令 $y_s=z_1=x_1$，此时 $r=3$，选第三行的基变量 w_3 出基，$y_s=x_1$ 进基，旋转变换后得表 11.4.4.

表 11.4.4　例 11.4.1 迭代三次的单纯形表

	w_1	w_2	w_3	x_1	x_2	u	z_0	
x_2	$\frac{1}{14}$	$-\frac{5}{28}$	$\frac{3}{28}$	0	1	$\frac{1}{7}$	0	$\frac{33}{14}$
z_0	$-\frac{1}{2}$	$-\frac{1}{4}$	$-\frac{1}{4}$	0	0	2^*	1	$\frac{3}{2}$
x_1	$-\frac{3}{14}$	$\frac{1}{28}$	$\frac{5}{28}^*$	1	0	$\frac{4}{7}$	0	$\frac{13}{14}$

令 $y_s = z_3 = u$，此时 $r=2$，选第二行的基变量 z_0 出基，$y_s = u$ 进基，旋转变换后得表 11.4.5.

表 11.4.5　例 11.4.1 迭代四次的单纯形表

	w_1	w_2	w_3	x_1	x_2	u	z_0	
x_2	$\frac{3}{28}$	$-\frac{9}{56}$	$\frac{7}{56}$	0	1	0	$-\frac{1}{14}$	$\frac{9}{4}$
u	$-\frac{1}{4}$	$-\frac{1}{8}$	$-\frac{1}{4}$	0	0	1	$\frac{1}{2}$	$\frac{3}{4}$
x_1	$-\frac{1}{14}$	$\frac{3}{28}$	$\frac{1}{4}^*$	1	0	0	$-\frac{2}{7}$	$\frac{1}{2}$

此时，z_0 为非基变量，$z_0 = 0$，即得到线性互补问题(11.4.7)的 CBF 解：

$$(w_1, w_2, w_3, x_1, x_2, u)^{\mathrm{T}} = \left(0, 0, 0, \frac{1}{2}, \frac{9}{4}, \frac{3}{4}\right)^{\mathrm{T}}.$$

故二次规划问题(11.4.6)的 K－T 点

$$(x_1, x_2)^{\mathrm{T}} = \left(\frac{1}{2}, \frac{9}{4}\right)^{\mathrm{T}}.$$

又问题(11.4.6)为凸规划，因此，其 K－T 点就是全局极小点.　□

从 Lemke 算法的迭代过程和本例可以看出，尽管 Lemke 算法很像线性规划的单纯形法，但它的换基原则实际上更为简单.

11.5　割平面法

前面几节主要讨论了二次规划的性质和算法，众所周知二次凸规划在二次规划中占有很重要的地位，很多算法对于二次凸规划更容易实现. 为此，本节将介绍一种常用的求解凸规划问题的方法——割平面法(cutting plane method).

割平面法是由 Kelley(1959)和 Cheney，Goidstenin(1959)独立提出的，也称

为 KCG 法. 其基本原理是将凸规划的可行解集合用有限个半空间来近似,并求解一系列不断改进的线性规划,它们的全局极小点收敛于原凸规划问题的全局极小点.

考虑凸规划问题

$$\begin{cases}\min & f(\boldsymbol{x});\\ \text{s.t.} & g_i(\boldsymbol{x})\geqslant 0,\quad i=1,2,\cdots,m,\end{cases}\tag{11.5.1}$$

其中 $f(\boldsymbol{x})$ 为 $\mathbb{R}^n$ 上可微的凸函数, $g_i(\boldsymbol{x})\ (i=1,\cdots,m)$ 是 $\mathbb{R}^n$ 上可微的凹函数,则(11.5.1)式等价于

$$\begin{cases}\min & z;\\ \text{s.t.} & z-f(\boldsymbol{x})\geqslant 0,\\ & g_i(\boldsymbol{x})\geqslant 0, i=1,2,\cdots,m.\end{cases}$$

这里 $z-f(\boldsymbol{x})$ 是 $\mathbb{R}^n$ 上的凹函数. 不失一般性,考虑问题

$$\begin{cases}\min & f(\boldsymbol{x})=\boldsymbol{c}^{\mathrm{T}}\boldsymbol{x};\\ \text{s.t.} & g_i(\boldsymbol{x})\geqslant 0, i=1,2,\cdots,m,\end{cases}\tag{11.5.2}$$

这里 $\boldsymbol{c}\in\mathbb{R}^n$, $g_i(\boldsymbol{x})\ (i=1,2,\cdots,m)$ 是可微的凹函数. 问题的可行域记作

$$S=\{\boldsymbol{x}\mid g_i(\boldsymbol{x})\geqslant 0, i=1,2,\cdots,m\},$$

根据例 2.4.3, S 为凸集.

割平面法的关键就是要用多胞形取代可行域 S,并在多胞形上极小化目标函数 $\boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}$. 因此首先给定一组线性约束,确定一个多胞形 S_1,使得 $S\subseteq S_1$,然后解线性规划问题

$$\begin{cases}\min & \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x};\\ \text{s.t.} & \boldsymbol{x}\in S_1.\end{cases}\tag{11.5.3}$$

设问题(11.5.3)的全局极小点为 $\boldsymbol{x}_1$,在 $\boldsymbol{x}_1$ 处将问题(11.5.2)中具有最小约束函数值的约束线性化,即当

$$g_r(\boldsymbol{x}_1)=\min\{g_i(\boldsymbol{x}_1)\mid i=1,2,\cdots,m\}$$

时,将约束 $g_r(\boldsymbol{x})\geqslant 0$ 线性化,并把线性化后的约束

$$g_r(\boldsymbol{x}_1)+\nabla g_r(\boldsymbol{x}_1)^{\mathrm{T}}(\boldsymbol{x}-\boldsymbol{x}_1)\geqslant 0$$

加入上述确定多胞形 S_1 的线性约束组中,构成一组新的线性约束,从而可确定一个新的多胞形 S_2,再解线性规划问题

$$\begin{cases}\min & \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x};\\ \text{s.t.} & \boldsymbol{x}\in S_2.\end{cases}\tag{11.5.4}$$

这时,必有 $S_2\subseteq S_1$.

另外,由于 $g_r(\boldsymbol{x})$ 是凹函数,因此

$$g_r(\boldsymbol{x})\leqslant g_r(\boldsymbol{x}_1)+\nabla g_r(\boldsymbol{x}_1)^{\mathrm{T}}(\boldsymbol{x}-\boldsymbol{x}_1),$$

从而满足约束 $g_r(\boldsymbol{x})\geqslant 0$ 的点 $\boldsymbol{x}$ 必使得

$$g_r(\boldsymbol{x}_1)+\nabla g_r(\boldsymbol{x}_1)^{\mathrm{T}}(\boldsymbol{x}-\boldsymbol{x}_1)\geqslant 0,$$

于是，$S\subseteq S_2$.

求得问题(11.5.4)的全局极小点 $\boldsymbol{x}_2$ 后，重复上述步骤，则可构造出一系列多胞形

$$S_k\subseteq S_{k-1}\subseteq\cdots\subseteq S_2\subseteq S_1,$$

且 $S\subseteq S_i(i=1,2,\cdots,k)$. 随着 k 不断增大，线性规划的可行域 S_k 将越来越逼近原凸规划问题的可行域 S，并且求解这一系列线性规划问题将产生全局极小点序列 $\{\boldsymbol{x}_k\}$，则序列 $\{\boldsymbol{x}_k\}$ 与凸规划问题(11.5.2)的全局极小点 $\bar{\boldsymbol{x}}$ 之间存在如下关系.

定理 11.5.1 设凸规划问题(11.5.2)满足

(1) $g_i(\boldsymbol{x})(i=1,2,\cdots,m)$ 为连续可微的凹函数；

(2) 迭代步骤中确定的一系列线性规划问题

$$\begin{cases}\min & \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x};\\ \text{s.t.} & \boldsymbol{x}\in S_k\end{cases}$$

的全局极小点 $\boldsymbol{x}_k$ 均存在，则有

(1) 若存在某个 k 使 $\boldsymbol{x}_k\in S$，则 $\boldsymbol{x}_k$ 即为问题(11.5.2)的全局极小点；

(2) 若情形(1)总不发生，则序列 $\{\boldsymbol{x}_k\}$ 的任意极限点都是问题(11.5.2)的全局极小点.

证明 (1) 若存在某个 k 使 $\boldsymbol{x}_k\in S$，又 $S\subseteq S_k$ 且 $\boldsymbol{x}_k$ 为 S_k 中目标函数 $\boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}$ 的全局极小点，故知 $\boldsymbol{x}_k$ 为 S 中 $\boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}$ 的全局极小点，即问题(11.5.2)的全局极小点.

(2) 由 S_k 的构造知，$S\subseteq S_{k+1}\subseteq S_k$，因此

$$\min_{x\in S}\boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}\geqslant\min_{x\in S_{k+1}}\boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}=\boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}_{k+1}\geqslant\min_{x\in S_k}\boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}=\boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}_k.$$

记

$$I=\{i\mid g_i(\boldsymbol{x}_k)\geqslant 0,1\leqslant i\leqslant m\},$$
$$\bar{I}=\{i\mid g_i(\boldsymbol{x}_k)<0,1\leqslant i\leqslant m\},$$

并设

$$\lim_{k\to\infty}\boldsymbol{x}_k=\tilde{\boldsymbol{x}},$$

由 $g_i(\boldsymbol{x})$ 的连续性知，必有

(i) $g_i(\tilde{\boldsymbol{x}})=\lim\limits_{k\to\infty}g_i(\boldsymbol{x}_k)\geqslant 0,i\in I$.

(ii) 当 $i\in\bar{I}$ 且 $r\geqslant 1$ 时，有

$$g_i(\boldsymbol{x}_k)+\nabla g_i(\boldsymbol{x}_k)^{\mathrm{T}}(\boldsymbol{x}_{k+r}-\boldsymbol{x}_k)\geqslant 0,$$

且

$$\tilde{\boldsymbol{x}}=\lim_{k\to\infty}\boldsymbol{x}_k=\lim_{k\to\infty}\boldsymbol{x}_{k+r}.$$

因此,对于任意 $i\in\bar{I}$,由 $g_i(\boldsymbol{x}_k)$ 的连续可微性即知,$\nabla g_i(\boldsymbol{x}_k)$ 有界,从而

$$g_i(\tilde{\boldsymbol{x}})=\lim_{k\to\infty}g_i(\boldsymbol{x}_k)=\lim_{k\to\infty}[g_i(\boldsymbol{x}_k)+\nabla g_i(\boldsymbol{x}_k)(\boldsymbol{x}_{k+r}-\boldsymbol{x}_k)]\geqslant 0,i\in\bar{I}.$$

综合(i)、(ii)可知,$\tilde{\boldsymbol{x}}\in S$.

又对于任意可行点 $\boldsymbol{x}\in S$,因 $S\subseteq S_k$,故 $\boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}\geqslant\boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}_k$,从而

$$\boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}\geqslant\lim_{k\to\infty}\boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}_k=\boldsymbol{c}^{\mathrm{T}}\tilde{\boldsymbol{x}}.$$

由此可知,$\tilde{\boldsymbol{x}}$ 是凸规划问题(11.5.2)的全局极小点. □

从计算角度来看,用割平面法求解一系列线性规划问题时,目标函数始终不变,线性约束每次增加一个,因此可用对偶单纯形法求解. 当约束每次增加一个时,只要在前一个线性规划问题的最优单纯形表上添加一行,经过几次旋转变换,很快就可获得新的线性规划问题的全局极小点. 割平面法具体步骤如下:

算法 11-3(割平面法)

Step 1 选取初始数据. 给定初始多胞形

$$S_1=\{\boldsymbol{x}\mid\boldsymbol{A}^{\mathrm{T}}\boldsymbol{x}-\boldsymbol{b}\geqslant\boldsymbol{0}\},$$

使得 $S\subseteq S_1$ 且目标函数 $\boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}$ 在 S_1 上有界,设允许误差为 $\varepsilon>0$,令 $k=1$.

Step 2 求解线性规划问题. 求解问题

$$\begin{cases}\min & \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x};\\ \text{s. t.} & \boldsymbol{x}\in S_k,\end{cases}$$

设其全局极小点为 $\boldsymbol{x}_k$.

Step 3 寻找 $\boldsymbol{x}_k$ 处的最小约束函数. 确定下标 $i_k(1\leqslant i_k\leqslant m)$,满足

$$g_{i_k}(\boldsymbol{x}_k)=\min_{1\leqslant i\leqslant m}g_i(\boldsymbol{x}_k),$$

若存在多个,任选其一.

Step 4 检验是否满足判别准则. 若

$$g_{i_k}(\boldsymbol{x}_k)>-\varepsilon,$$

则停止计算,$\boldsymbol{x}_k$ 即为原凸规划问题的近似全局极小点;否则,令

$$S_{k+1}=S_k\cap\{\boldsymbol{x}\mid g_{i_k}(\boldsymbol{x}_k)+\nabla g_{i_k}(\boldsymbol{x}_k)^{\mathrm{T}}(\boldsymbol{x}-\boldsymbol{x}_k)\geqslant 0\},$$

置 $k:=k+1$,返回 Step 2.

值得注意的是,算法的 Step1 中要选取包含 S 在内的初始多胞形 S_1,一般可采用以下两种方法:

(1) 选取较简单的多胞形 S_1(比如充分大的超立方体)包含 S,使所得线性规划问题有解且易求解;

(2) 若对可行域 S 的大小难以估计,可先在 S 中任选一点 $\boldsymbol{x}_0$,在 $\boldsymbol{x}_0$ 处对所有 $g_i(\boldsymbol{x})$ 利用 Taylor 展开式可得

$$g_i(\boldsymbol{x})\approx g_i(\boldsymbol{x}_0)+\nabla g_i(\boldsymbol{x}_0)^{\mathrm{T}}(\boldsymbol{x}-\boldsymbol{x}_0),i=1,2,\cdots,m.$$

再取

$$S_1=\{\boldsymbol{x}\mid g_i(\boldsymbol{x}_0)+\nabla g_i(\boldsymbol{x}_0)^{\mathrm{T}}(\boldsymbol{x}-\boldsymbol{x}_0)\geqslant 0, i=1,2,\cdots,m\},$$

易证 $S\subseteq S_1$.

例 11.5.1 求解凸规划问题

$$\begin{cases}\min & f(\boldsymbol{x})=2x_1+x_2;\\ \text{s.t.} & 4-(x_1-3)^2-(x_2-2)^2\geqslant 0,\\ & x_2-(x_1-3)^2\geqslant 0.\end{cases} \tag{11.5.5}$$

解 由 $g_1(\boldsymbol{x})\geqslant 0$ 知, $|x_1-3|\leqslant 2$, $|x_2-2|\leqslant 2$, 即 $1\leqslant x_1\leqslant 5, 0\leqslant x_2\leqslant 4$ 故取

$$S_1=\{(x_1,x_2)^{\mathrm{T}}\mid 0\leqslant x_1\leqslant 5, 0\leqslant x_2\leqslant 4\}.$$

显然, $S\subseteq S_1$.

解线性规划问题

$$\begin{cases}\min & f(\boldsymbol{x})=2x_1+x_2;\\ \text{s.t.} & 0\leqslant x_1\leqslant 5,\\ & 0\leqslant x_2\leqslant 4.\end{cases}$$

用单纯形法或图解法可求得其全局极小点为 $\boldsymbol{x}_1=(0,0)^{\mathrm{T}}$.

由于 $g_1(\boldsymbol{x}_1)=-9, g_2(\boldsymbol{x}_1)=-9$, 因此任取其一, 令 $i_1=2$, 并将 $g_2(\boldsymbol{x})$ 在点 $\boldsymbol{x}_1$ 处线性化, 即

$$g_2(\boldsymbol{x})\approx 6x_1+x_2-9.$$

令 $S_2=S_1\cap\{(x_1,x_2)^{\mathrm{T}}\mid 6x_1+x_2-9\geqslant 0\}$, 解线性规划问题

$$\begin{cases}\min & f(\boldsymbol{x})=2x_1+x_2;\\ \text{s.t.} & \boldsymbol{x}\in S_2.\end{cases}$$

用对偶单纯形法求得其全局极小点为 $\boldsymbol{x}_2=\left(\dfrac{3}{2},0\right)^{\mathrm{T}}$.

因 $g_1(\boldsymbol{x}_2)=-\dfrac{9}{4}, g_2(\boldsymbol{x}_2)=-\dfrac{9}{4}$, 故仍取 $i_2=2$, 并将 $g_2(\boldsymbol{x})$ 在点 $\boldsymbol{x}_2$ 处线性化, 即

$$g_2(\boldsymbol{x})\approx 3x_1+x_2-\frac{27}{4}.$$

令 $S_3=S_2\cap\left\{(x_1,x_2)^{\mathrm{T}}\,\middle|\, 3x_1+x_2-\dfrac{27}{4}\geqslant 0\right\}$, 解线性规划问题

$$\begin{cases}\min & f(\boldsymbol{x})=2x_1+x_2;\\ \text{s.t.} & \boldsymbol{x}\in S_3.\end{cases}$$

用对偶单纯形法求得其全局极小点为 $\boldsymbol{x}_3=\left(\dfrac{9}{4},0\right)^{\mathrm{T}}$.

由于 $g_1(\boldsymbol{x}_3)=-\dfrac{9}{16}, g_2(\boldsymbol{x}_3)=-\dfrac{9}{16}$，因此仍取 $i_3=2$，并将 $g_2(\boldsymbol{x})$ 在点 $\boldsymbol{x}_3$ 处线性化，即

$$g_2(\boldsymbol{x})\approx\frac{3}{2}x_1+x_2-\frac{63}{16}.$$

令 $S_4=S_3\cap\left\{(x_1,x_2)^{\mathrm{T}}\,\middle|\,\dfrac{3}{2}x_1+x_2-\dfrac{63}{16}\geqslant 0\right\}$，解线性规划问题

$$\begin{cases}\min & f(\boldsymbol{x})=2x_1+x_2;\\ \text{s.t.} & \boldsymbol{x}\in S_4,\end{cases}$$

其全局极小点为 $\boldsymbol{x}_4=\left(\dfrac{15}{8},\dfrac{9}{8}\right)^{\mathrm{T}}$. 此时 $\boldsymbol{x}_4$ 已接近问题(11.5.4)的全局极小点 $\bar{\boldsymbol{x}}=(2,1)^{\mathrm{T}}$，$f(\boldsymbol{x}_4)=\dfrac{39}{8}$ 也接近目标函数的最优值 $f(\bar{\boldsymbol{x}})=5$. 若需得到更精确的近似全局极小点，可仿照上述方法继续迭代下去. □

割平面法是求解凸规划问题最早的方法之一，由于简单易行、比较直观，因此在实际工作中得到了广泛使用. 但在迭代中，当 k 充分大时，势必得到一个庞大的线性不等式约束组，这给相应的线性规划问题的求解带来了困难.

习 题 十 一

1. 求下述二次规划问题的全局极小点 $\bar{\boldsymbol{x}}$ 及相应的 Lagrange 乘子 $\bar{v}$:

$$\begin{cases}\min & x_1^2+4x_2^2+5x_3^2;\\ \text{s.t.} & 3x_1+4x_2=13.\end{cases}$$

2. 用直接消去法和广义消去法求解下列问题:

(1) $\begin{cases}\min & 2x_1^2+x_2^2+x_1x_2-x_1-x_2;\\ \text{s.t.} & x_1+x_2=1.\end{cases}$

(2) $\begin{cases}\min & \dfrac{3}{2}x_1^2-x_1x_2+x_2^2-x_2x_3+\dfrac{1}{2}x_3^2+x_1+x_2+x_3;\\ \text{s.t.} & x_1+2x_2+x_3=4.\end{cases}$

3. 用起作用集法求解二次凸规划问题(取初始可行点 $\boldsymbol{x}_1=(0,0)^{\mathrm{T}}$):

(1) $\begin{cases}\min & 9x_1^2+9x_2^2-30x_1-72x_2;\\ \text{s.t.} & -2x_1+x_2\geqslant -4,\\ & x_1\geqslant 0, x_2\geqslant 0.\end{cases}$

(2) $\begin{cases}\min & x_1^2-x_1x_2+x_2^2-3x_1;\\ \text{s.t.} & -x_1-x_2\geqslant -2,\\ & x_1\geqslant 0, x_2\geqslant 0.\end{cases}$

4. 用 Wolfe 算法求解问题:

$$\begin{cases} \min & f(\boldsymbol{x}) = x_1^2 + x_2^2 + x_3^2 - 2x_1x_2; \\ \text{s.t.} & 2x_1 + x_2 - 4 = 0, \\ & 5x_2 - x_3 - 7 = 0, \\ & x_1 \geqslant 0, x_2 \geqslant 0, x_3 \geqslant 0. \end{cases}$$

5. 用 Lemke 算法求解下列问题：

(1) $$\begin{cases} \min & 2x_1^2 + x_2^2 - 2x_1x_2 - 6x_1 - 2x_2; \\ \text{s.t.} & -x_1 - x_2 + 2 \geqslant 0, \\ & -2x_1 + x_2 + 2 \geqslant 0, \\ & x_1 \geqslant 0, x_2 \geqslant 0. \end{cases}$$

(2) $$\begin{cases} \min & 2x_1^2 + 2x_2^2 + x_3^2 + 2x_1x_2 + 2x_1x_3 - 8x_1 - 6x_2; \\ \text{s.t.} & -x_1 - x_2 - x_3 + 3 \geqslant 0, \\ & x_1 \geqslant 0, x_2 \geqslant 0, x_3 \geqslant 0. \end{cases}$$

6. 对于凸规划问题

$$\begin{cases} \min & x_1 - x_2; \\ \text{s.t.} & 25 - x_1^2 - x_2^2 \geqslant 0, \\ & 2x_1 + x_2 - 1 \geqslant 0, \end{cases}$$

从 $S_1 = \{(x_1, x_2) \mid -5 \leqslant x_1 \leqslant 5, -5 \leqslant x_2 \leqslant 5\}$ 出发，利用割平面法求迭代点列中的 $\boldsymbol{x}_1, \boldsymbol{x}_2$ 和 $\boldsymbol{x}_3$.

第十二章　线性分式规划

线性分式规划是非线性规划的一个特殊类型，其目标函数是两个线性函数之比，而约束是线性约束. 正如在第一章所看到的那样，它有着广泛的实际背景. 在这一章里，我们要针对线性分式规划(简记为 LFP)的特点，建立求解(LFP)的一些特别的算法.

12.1　原始单纯形法

考虑线性分式规划(LFP)问题

$$\begin{cases} \min \quad f(\boldsymbol{x}) = \dfrac{\boldsymbol{p}^{\mathrm{T}}\boldsymbol{x} + \alpha}{\boldsymbol{q}^{\mathrm{T}}\boldsymbol{x} + \beta}; \\ \text{s.t.} \quad \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}, \\ \qquad\quad \boldsymbol{x} \geqslant \boldsymbol{0}, \end{cases} \tag{12.1.1}$$

其中 $\boldsymbol{A} \in \mathbb{R}^{m\times n}, \boldsymbol{b} \in \mathbb{R}^m, \boldsymbol{p}, \boldsymbol{q}, \boldsymbol{x} \in \mathbb{R}^n, \alpha, \beta \in \mathbb{R}, \operatorname{rank} \boldsymbol{A} = m$. 记 $S = \{\boldsymbol{x} \mid \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}, \boldsymbol{x} \geqslant \boldsymbol{0}\}$ 为问题(12.1.1)的可行解集，假定

(1) $\boldsymbol{q}^{\mathrm{T}}\boldsymbol{x} + \beta > 0, \forall \boldsymbol{x} \in S$;

(2) S 是 $\mathbb{R}^n$ 中的非空有界集，

称满足假设条件(1)、(2)的 LFP 为简单问题. 易知，LFP 简单问题必有最优解.

本章只研究 LFP 简单问题.

定理 12.1.1　LFP 在多面体 S 的极点上取到最优解.

证明　设 $\boldsymbol{y}_1, \boldsymbol{y}_2, \cdots, \boldsymbol{y}_s$ 为 S 的全部极点，由多面体表示定理(即推论 2.2.6)可知，对任意 $\boldsymbol{x} \in S$，有

$$\boldsymbol{x} = \sum_{i=1}^{s} \alpha_i \boldsymbol{y}_i,$$

其中 $\alpha_i \geqslant 0, i = 1, 2, \cdots, s,\ \sum\limits_{i=1}^{s} \alpha_i = 1$.

记

$$u(\boldsymbol{x}) = \boldsymbol{p}^{\mathrm{T}}\boldsymbol{x} + \alpha, v(\boldsymbol{x}) = \boldsymbol{q}^{\mathrm{T}}\boldsymbol{x} + \beta, \tag{12.1.2}$$

$$f(\boldsymbol{y}_k) = \min\{f(\boldsymbol{y}_i) \mid i = 1, 2, \cdots, s\},$$

则有

$$u(\boldsymbol{y}_k)v(\boldsymbol{y}_i)\leqslant v(\boldsymbol{y}_k)u(\boldsymbol{y}_i),i=1,2,\cdots,s.$$

上式两边同乘以 $\alpha_i\geqslant 0,i=1,2,\cdots,s$,并把它们相加得

$$u(\boldsymbol{y}_k)\sum_{i=1}^{s}\alpha_i v(\boldsymbol{y}_i)\leqslant v(\boldsymbol{y}_k)\sum_{i=1}^{s}\alpha_i u(\boldsymbol{y}_i). \tag{12.1.3}$$

注意到 $u(\boldsymbol{y}_i)=\boldsymbol{p}^{\mathrm{T}}\boldsymbol{y}_i+\alpha,v(\boldsymbol{y}_i)=\boldsymbol{q}^{\mathrm{T}}\boldsymbol{y}_i+\beta,i=1,2,\cdots,s$,则有

$$\sum_{i=1}^{s}\alpha_i u(\boldsymbol{y}_i)=\boldsymbol{p}^{\mathrm{T}}(\sum_{i=1}^{s}\alpha_i\boldsymbol{y}_i)+\alpha(\sum_{i=1}^{s}\alpha_i)=\boldsymbol{p}^{\mathrm{T}}\boldsymbol{x}+\alpha=u(\boldsymbol{x}),$$

以及

$$\sum_{i=1}^{s}\alpha_i v(\boldsymbol{y}_i)=\boldsymbol{q}^{\mathrm{T}}\boldsymbol{x}+\beta=v(\boldsymbol{x}).$$

将以上两式代入(12.1.3)式得

$$u(\boldsymbol{y}_k)v(\boldsymbol{x})\leqslant v(\boldsymbol{y}_k)u(\boldsymbol{x}),$$

即有

$$\frac{u(\boldsymbol{y}_k)}{v(\boldsymbol{y}_k)}\leqslant\frac{u(\boldsymbol{x})}{v(\boldsymbol{x})},$$

亦即

$$f(\boldsymbol{y}_k)\leqslant f(\boldsymbol{x}),\forall\boldsymbol{x}\in S,$$

因此,$\boldsymbol{y}_k$ 是问题(12.1.1)的最优解. □

根据这一定理,可以从 LFP 的可行域的极点上寻找 LFP 的最优解.这样,我们可以从一个极点出发,利用线性规划单纯形迭代程序移动到另一个极点,最后求得最优解.

设 $\boldsymbol{B}$ 为一可行基,$\boldsymbol{A}=(\boldsymbol{B},\boldsymbol{N})$,则将

$$\boldsymbol{x}_B=\boldsymbol{B}^{-1}\boldsymbol{b}-\boldsymbol{B}^{-1}\boldsymbol{N}\boldsymbol{x}_N \tag{12.1.4}$$

代入(12.1.2)式得

$$\begin{aligned}u(\boldsymbol{x})&=\boldsymbol{p}_B^{\mathrm{T}}\boldsymbol{x}_B+\boldsymbol{p}_N^{\mathrm{T}}\boldsymbol{x}_N+\alpha\\&=(\boldsymbol{p}_B^{\mathrm{T}}\boldsymbol{B}^{-1}\boldsymbol{b}+\alpha)-(\boldsymbol{p}_B^{\mathrm{T}}\boldsymbol{B}^{-1}\boldsymbol{N}-\boldsymbol{p}_N^{\mathrm{T}})\boldsymbol{x}_N, \quad (12.1.5)\\v(\boldsymbol{x})&=\boldsymbol{q}_B^{\mathrm{T}}\boldsymbol{x}_B+\boldsymbol{p}_N^{\mathrm{T}}\boldsymbol{x}_N+\beta\\&=(\boldsymbol{q}_B^{\mathrm{T}}\boldsymbol{B}^{-1}\boldsymbol{b}+\beta)-(\boldsymbol{q}_B^{\mathrm{T}}\boldsymbol{B}^{-1}\boldsymbol{N}-\boldsymbol{q}_N^{\mathrm{T}})\boldsymbol{x}_N. \quad (12.1.6)\end{aligned}$$

记 $\boldsymbol{x}_B=(x_{j_1},x_{j_2},\cdots,x_{j_m})^{\mathrm{T}}$,用 R 表示非基变量指标集,把(12.1.4)、(12.1.5)和(12.1.6)三式相应地写成

$$x_{j_i}+\sum_{j\in R}b_{ij}x_j=b_{i0},i=1,2,\cdots,m, \tag{12.1.4a}$$

$$u+\sum_{j\in R}u_jx_j=u_0, \tag{12.1.5a}$$

$$v+\sum_{j\in R}v_jx_j=v_0, \tag{12.1.6a}$$

称

$$\begin{cases} \min \quad f = \dfrac{u_0 - \sum\limits_{j \in R} u_j x_j}{v_0 - \sum\limits_{j \in R} v_j x_j}; \\ \text{s.t.} \quad x_{j_i} + \sum\limits_{j \in R} b_{ij} x_j = b_{i0}, i = 1,2,\cdots,m, \\ \qquad\quad x_j \geqslant 0, j = 1,2,\cdots,n \end{cases}$$

为 LFP 的典式. 记

$$\pi_j = \begin{cases} v_0 u_j - u_0 v_j, & j \in R, \\ 0, & j \in \{1,2,\cdots,n\} \setminus R, \end{cases} \tag{12.1.7}$$

我们得到如下最优性定理.

定理 12.1.2 对可行基 $\boldsymbol{B}$,若所有 $\pi_j \leqslant 0, j \in R$,则对应的基本可行解 $\bar{\boldsymbol{x}}$:

$$\bar{x}_{j_i} = b_{i0}(i = 1,2,\cdots,m), \bar{x}_j = 0(j \in R)$$

为(LFP)的最优解.

证明 对任意 $\boldsymbol{x} \in S$,有

$$\begin{aligned} v_0 u(\boldsymbol{x}) - u_0 v(\boldsymbol{x}) &= v_0(u_0 - \sum_{j \in R} u_j x_j) - u_0(v_0 - \sum_{j \in R} v_j x_j) \\ &= -\sum_{j \in R} (v_0 u_j - u_0 v_j) x_j = -\sum_{j \in R} \pi_j x_j. \end{aligned}$$

若 $\pi_j \leqslant 0, j \in R$,则由上式得

$$v_0 u(\boldsymbol{x}) - u_0 v(\boldsymbol{x}) \geqslant 0,$$

即有

$$\frac{u(\boldsymbol{x})}{v(\boldsymbol{x})} \geqslant \frac{u(\bar{\boldsymbol{x}})}{v(\bar{\boldsymbol{x}})} = \frac{u_0}{v_0},$$

即

$$f(\boldsymbol{x}) \geqslant f(\bar{\boldsymbol{x}}), \forall \boldsymbol{x} \in S,$$

所以 $\bar{\boldsymbol{x}}$ 为 LFP 的最优解. □

称 π_j 为变量 x_j 关于可行基 $\boldsymbol{B}$ 的检验数. 称

$$\pi_j \leqslant 0, j = 1,2,\cdots,n$$

为 LFP 的最优性条件.

当某个 $\pi_k > 0$ 时,可以借助于类似于线性规划的原始单纯形法来修改基本可行解. 让 x_k 进基,由

$$\frac{b_{r0}}{b_{rk}} = \theta = \min\left\{\frac{b_{i0}}{b_{ik}} \,\middle|\, b_{ik} > 0, i = 1,2,\cdots,m\right\} \tag{12.1.8}$$

确定 x_{j_r}为离基变量. 从

$$x_{j_r} + \sum_{j \in R} b_{rj} x_j = b_{r0}$$

式中解出 x_k，并代入(12.1.4a)、(12.1.5a)和(12.1.6a)式中得

$$x_k + \sum_{j\in R\setminus\{k\}} \frac{b_{rj}}{b_{rk}} x_j + \frac{1}{b_{rk}} x_{j_r} = \frac{b_{r0}}{b_{rk}},$$

$$x_{j_i} + \sum_{j\in R\setminus\{k\}} \left(b_{ij} - \frac{b_{rj}}{b_{rk}} b_{ik}\right) x_j - \frac{1}{b_{rk}} b_{ik} x_{j_r} = b_{i0} - \frac{b_{r0}}{b_{rk}} b_{ik}, i \in I\setminus\{r\},$$

$$u + \sum_{j\in R\setminus\{k\}} \left(u_j - \frac{b_{rj}}{b_{rk}} u_k\right) x_j - \frac{1}{b_{rk}} u_k x_{j_r} = u_0 - \frac{b_{r0}}{b_{rk}} u_k,$$

$$v + \sum_{j\in R\setminus\{k\}} \left(v_j - \frac{b_{rj}}{b_{rk}} v_k\right) x_j - \frac{1}{b_{rk}} v_k x_{j_r} = v_0 - \frac{b_{r0}}{b_{rk}} v_k,$$

对应的基本可行解 $\hat{\boldsymbol{x}}$ 为

$$\hat{x}_k = \frac{b_{r0}}{b_{rk}} = \theta, \hat{x}_{j_i} = b_{i0} - \frac{b_{r0}}{b_{rk}} b_{ik} = b_{i0} - \theta b_{ik}, i \in \{1,2,\cdots,m\}\setminus\{r\},$$

$$\hat{x}_j = 0, j \in \hat{R} = (R\setminus\{k\}) \cup \{j_r\}, \tag{12.1.9}$$

对应的目标函数值为

$$\hat{f} = \frac{u_0 - \theta u_k}{v_0 - \theta v_k} = \frac{\hat{u}_0}{\hat{v}_0}. \tag{12.1.10}$$

这一迭代步骤可在类似于线性规划的单纯形表 12.1.1 上进行. 与线性规划单纯形表不同的是，这里按(12.1.5a)、(12.1.6a)添加了两行，称为目标行. 基变量所在行为约束行，最后一行为检验行. 迭代时，以 b_{rk} 为主元，对约束行和目标行作旋转变换，但检验行中的检验数要按(12.1.7)式计算.

表 12.1.1　原始单纯形法迭代表格样式

	x_{j_1}	$\cdots$	x_{j_m}	$x_j(j\in R)$	
x_{j_1}	1	$\cdots$	0	b_{1j}	b_{10}
$\vdots$	$\vdots$		$\vdots$	$\vdots$	$\vdots$
x_{j_m}	0	$\cdots$	1	b_{mj}	b_{m0}
u	0	$\cdots$	0	u_j	u_0
v	0	$\cdots$	0	v_j	v_0
f	0	$\cdots$	0	$\pi_j = v_0 u_j - u_0 v_j$	u_0/v_0

下面我们证明，当 $\pi_k > 0$ 时，通过一次迭代到达另一极点，目标数值有所下降.

引理 12.1.3　当 $\pi_k > 0$ 时，至少存在一个 $r \in \{1,2,\cdots,m\}$，使得

$$b_{rk} > 0.$$

证明　用反证法. 若不然，对一切 $i \in \{1,2,\cdots,m\}$，都有 $b_{ik} \leqslant 0$，那么，任取数 $\delta > 0$，按(12.1.4a)作一区别于基本可行解 $\boldsymbol{x}$ 的新的可行解 $\tilde{\boldsymbol{x}}$:

$$\tilde{x}_k = \delta > 0,$$
$$\tilde{x}_j = 0, j \in R \setminus \{k\},$$
$$\tilde{x}_{j_i} = b_{i0} - b_{ik} \geqslant 0, i = 1,2,\cdots,m.$$

由于对任意 $\delta > 0$，$\tilde{x}$ 都是 LFP 的可行解，这与假设 S 为有界域矛盾. □

引理 12.1.3 保证，如果 $\pi_k > 0$，选取 x_k 为进基变量，那么一定可以根据(12.1.8)式来选出离基变量 x_{j_r}.

定理 12.1.4 由一可行基开始，每经过一次单纯形迭代都得到一个基本可行解，且目标函数值单调递减.

证明 由(12.1.8)及(12.1.9)式易知 $\hat{\boldsymbol{x}} \geqslant \boldsymbol{0}$，故经过一次单纯形迭代后仍然得到基本可行解. 由(12.1.10)式知

$$\begin{aligned} v_0 \hat{u}_0 - u_0 \hat{v}_0 &= v_0 (u_0 - \theta u_k) - u_0 (v_0 - \theta v_k) \\ &= -\theta (v_0 u_k - u_0 v_k) = -\theta \pi_k \leqslant 0, \end{aligned}$$

故 $\dfrac{u_0}{v_0} \geqslant \dfrac{\hat{u}_0}{\hat{v}_0}$，即目标函数值单调递减. 特别，在非退化情况下，$\theta > 0$，目标函数严格单调递减. □

算法 12-1（原始单纯形法）

Step 1 给定一个初始可行基 $\boldsymbol{B}$，按表 12.1.1 的格式写出相应的单纯形表 $T(\boldsymbol{B})$.

Step 2 如果所有 $\pi_j \leqslant 0, j = 1,2,\cdots,n$，则关于 $\boldsymbol{B}$ 的可行基 $\bar{\boldsymbol{x}}$ 为最优解，停止计算；否则，转 Step 3.

Step 3 确定进基变量. 令 $\pi_k = \max\{\pi_j \mid j \in R\}$，以 x_k 为进基变量.

Step 4 确定离基变量. 由

$$\frac{b_{r0}}{b_{rk}} = \theta = \min\left\{\frac{b_{i0}}{b_{ik}} \,\middle|\, b_{ik} > 0, i = 1,2,\cdots,m\right\}$$

确定 x_{j_r} 为离基变量.

Step 5 以 b_{rk} 为主元对约束行和目标行作旋转变换，并计算各检验数，得新的单纯形表 $T(\widehat{\boldsymbol{B}})$，以 $\widehat{\boldsymbol{B}}$ 代替 $\boldsymbol{B}$，转 Step 2.

在非退化情况下，算法从一个极点移到另一个极点，目标函数每次迭代严格单调递减，因而每次产生的极点是不同的. 由于仅仅存在有限个极点，因此算法在有限步停止.

在退化情况下，算法可能会出现“基循环”，从而不能在有限步结束. 文献[10]给出了这样的例子并证明了引入 Bland 规则可使算法在有限步结束.

例 12.1.1 用线性分式规划原始单纯形法求解下列问题：

$$\begin{cases} \min & \dfrac{-2x_1 + x_2 + 2}{x_1 + 3x_2 + 4}; \\ \text{s.t.} & -x_1 + x_2 \leqslant 4, \\ & x_2 \leqslant 6, \\ & 2x_1 + x_2 \leqslant 14, \\ & x_1, x_2 \geqslant 0. \end{cases} \qquad (12.1.11)$$

解 引入松弛变量 x_3,x_4 和 x_5，我们得到初始基本可行解 $x_0=(0,0,4,6,14)^T$. 初始单纯形表如表 12.1.2 所示.

表 12.1.2 原始单纯形法初始单纯形表

	x_1	x_2	x_3	x_4	x_5	
x_3	-1	1	1	0	0	4
x_4	0	1	0	1	0	6
x_5	2^*	1	0	0	1	14
u	2	-1	0	0	0	2
v	-1	-3	0	0	0	4
f	10	2	0	0	0	$\frac{1}{2}$

由表 12.1.2 知，$\pi_1=\max\{\pi_j\mid j\in R\}$，以 x_1 为进基变量，并由 $\theta=\min\left\{\frac{14}{2}\right\}=\frac{b_{50}}{b_{51}}$确定 x_5 为离基变量. 以 $b_{51}=2$ 为主元进行迭代，迭代结果如表 12.1.3 所示.

在表 12.1.3 中，$\pi_j\leqslant 0, j=1,2,\cdots,n$，则表 12.1.3 已是最优表. 从而得到问题 (12.1.11) 的最优解为 $\bar{x}=(7,0,11,6,0)^T$，对应的目标函数最优值为 $\bar{f}=-\frac{12}{11}$. □

表 12.1.3 原始单纯形法最优单纯形表

	x_1	x_2	x_3	x_4	x_5	
x_3	0	$\frac{3}{2}$	1	0	$\frac{1}{2}$	11
x_4	0	1	0	1	0	6
x_1	1	$\frac{1}{2}$	0	0	$\frac{1}{2}$	7
u	0	-2	0	0	-1	-12
v	0	-2	0	0	$\frac{1}{2}$	11
f	0	-46	0	0	-5	$-\frac{12}{11}$

12.2 Gilmore-Gomory 方法

Gilmore 和 Gomory 利用既约梯度于 1963 年提出了一种求解 LFP 的单纯形

法. 可以证明,上一节讲述的原始单纯形法是 Gilmore 和 Gomory 法的简化.

定理 12.2.1 设 $\bar{\boldsymbol{x}}$ 是问题(12.1.1)的 K - T 点,则 $\bar{\boldsymbol{x}}$ 为全局极小点.

证明 对任意 $\boldsymbol{x}\in S$,有

$$\nabla f(\boldsymbol{x})=\frac{(\boldsymbol{q}^{\mathrm{T}}\boldsymbol{x}+\beta)\boldsymbol{p}-(\boldsymbol{p}^{\mathrm{T}}\boldsymbol{x}+\alpha)\boldsymbol{q}}{(\boldsymbol{q}^{\mathrm{T}}\boldsymbol{x}+\beta)^{2}}. \tag{12.2.1}$$

因为 $\bar{\boldsymbol{x}}$ 是 K - T 点,所以它满足下述 K - T 条件:存在向量 $\boldsymbol{u}\geqslant\boldsymbol{0}$ 和向量 $\boldsymbol{v}$,使得

$$\begin{gathered}\nabla f(\bar{\boldsymbol{x}})-\boldsymbol{A}^{\mathrm{T}}\boldsymbol{v}-\boldsymbol{u}=\boldsymbol{0},\\ \boldsymbol{u}^{\mathrm{T}}\bar{\boldsymbol{x}}=0,\end{gathered} \tag{12.2.2}$$

从而 $\forall\,\boldsymbol{x}\in S$,有

$$\nabla f(\bar{\boldsymbol{x}})^{\mathrm{T}}(\boldsymbol{x}-\bar{\boldsymbol{x}})-\boldsymbol{v}^{\mathrm{T}}\boldsymbol{A}(\boldsymbol{x}-\bar{\boldsymbol{x}})-\boldsymbol{u}^{\mathrm{T}}(\boldsymbol{x}-\bar{\boldsymbol{x}})=0.$$

注意到 $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{b},\boldsymbol{x}\geqslant\boldsymbol{0}$ 及 $\boldsymbol{A}\bar{\boldsymbol{x}}=\boldsymbol{b}$,由(12.2.2)式及 $\boldsymbol{u}\geqslant\boldsymbol{0}$ 得

$$\nabla f(\bar{\boldsymbol{x}})^{\mathrm{T}}(\boldsymbol{x}-\bar{\boldsymbol{x}})=\boldsymbol{u}^{\mathrm{T}}\boldsymbol{x}\geqslant 0. \tag{12.2.3}$$

将(12.2.1)式代入(12.2.3)式并整理得

$$(\boldsymbol{q}^{\mathrm{T}}\bar{\boldsymbol{x}}+\beta)(\boldsymbol{p}^{\mathrm{T}}\boldsymbol{x}+\alpha)\geqslant(\boldsymbol{q}^{\mathrm{T}}\boldsymbol{x}+\beta)(\boldsymbol{p}^{\mathrm{T}}\bar{\boldsymbol{x}}+\alpha).$$

对于 LFP 简单问题有

$$\frac{\boldsymbol{p}^{\mathrm{T}}\boldsymbol{x}+\alpha}{\boldsymbol{q}^{\mathrm{T}}\boldsymbol{x}+\beta}\geqslant\frac{\boldsymbol{p}^{\mathrm{T}}\bar{\boldsymbol{x}}+\alpha}{\boldsymbol{q}^{\mathrm{T}}\bar{\boldsymbol{x}}+\beta}.$$

即 $f(\boldsymbol{x})\geqslant f(\bar{\boldsymbol{x}})$, $\forall\,\boldsymbol{x}\in S$,因此 $\bar{\boldsymbol{x}}$ 为全局极小点. □

在非退化假设下,设 $\boldsymbol{B}$ 是问题(12.1.1)的一个可行基,$\boldsymbol{x}=\begin{bmatrix}\boldsymbol{x}_B\\ \boldsymbol{x}_N\end{bmatrix}$ 是一个非退化的基本可行解,由(12.1.5)和(12.1.6)式容易得到 $f(\boldsymbol{x})$ 的既约梯度由下式给出:

$$\boldsymbol{r}(\boldsymbol{x}_N)^{\mathrm{T}}=\frac{-(\boldsymbol{p}_B^{\mathrm{T}}\boldsymbol{B}^{-1}\boldsymbol{N}-\boldsymbol{p}_N^{\mathrm{T}})(\boldsymbol{q}^{\mathrm{T}}\boldsymbol{x}+\beta)+(\boldsymbol{q}_B^{\mathrm{T}}\boldsymbol{B}^{-1}\boldsymbol{N}-\boldsymbol{q}_N^{\mathrm{T}})(\boldsymbol{p}^{\mathrm{T}}\boldsymbol{x}+\alpha)}{(\boldsymbol{q}^{\mathrm{T}}\boldsymbol{x}+\beta)^{2}}$$

由(9.3.11)式并注意 $\boldsymbol{x}_N=\boldsymbol{0}$ 知,问题(12.1.1)的 K - T 条件等价于

$$\boldsymbol{r}(\boldsymbol{x}_N)\geqslant\boldsymbol{0}. \tag{12.2.4}$$

根据定理 12.2.1,(12.2.4)式为问题(12.1.1)的最优性充分条件.

当 $\boldsymbol{r}(\boldsymbol{x}_N)\not\geqslant\boldsymbol{0}$ 时,我们利用单纯形迭代程序从一个极点移到另一个极点,并使目标函数值有所下降. 令

$$-r_k=\max\{-r_j\mid r_j<0,j\in R\},$$

其中 r_j 是 $\boldsymbol{r}(\boldsymbol{x}_N)$ 中对应 x_j 的分量. 以 x_k 为进基变量. 由

$$\frac{b_{r0}}{b_{rk}}=\theta=\min\left\{\frac{b_{i0}}{b_{ik}}\;\middle|\;b_{ik}>0,i=1,2,\cdots,m\right\} \tag{12.2.5}$$

确定离基变量 x_{j_r}. 正如引理 12.1.3 指出的那样,如果 $\boldsymbol{r}(\boldsymbol{x}_N)\not\geqslant\boldsymbol{0}$,那么对于 LFP 简单问题,总可由(12.2.5)式选出离基变量. 这样,经过 $\{k,r\}$ 旋转变换,将迭代点从一个极点移到了另一个极点. 此时,搜索方向 $\boldsymbol{d}=\begin{bmatrix}\boldsymbol{d}_B\\ \boldsymbol{d}_N\end{bmatrix}$ 定义为

$$\boldsymbol{d}_B=-\boldsymbol{B}^{-1}\boldsymbol{N}\boldsymbol{d}_N,\quad d_{N_j}=\begin{cases}1, & j=k,\\ 0, & j\in R\setminus\{k\}.\end{cases}$$

容易验证 $\boldsymbol{d}$ 满足(9.3.6)和(9.3.8)式,因此 $\boldsymbol{d}$ 是一个可行下降方向. 这样我们

就证明了在非退化情况下，按上述单纯形迭代法，迭代点从一个极点移到另一个极点时，目标函数值严格单调递减．进一步，算法将在有限步结束．

算法 12－2（Gilmore 和 Gomory 方法）

Step 1　给定一个初始可行基 $\boldsymbol{B}$，按表 12.2.1 的格式写出相应的单纯形表 $T(\boldsymbol{B})$．

Step 2　计算 $\boldsymbol{r}(\boldsymbol{x}_N)$，如果 $\boldsymbol{r}(\boldsymbol{x}_N)\geqslant\boldsymbol{0}$，则关于 $\boldsymbol{B}$ 的基本可行解 $\bar{\boldsymbol{x}}$ 为最优解，停止计算；否则，转 Step 3.

Step 3　确定进基变量．令

$$-r_k=\max\{-r_j\mid r_j<0,j\in R\},$$

以 x_k 为进基变量．

Step 4　确定离基变量．由

$$\frac{b_{r0}}{b_{rk}}=\theta=\min\left\{\frac{b_{i0}}{b_{ik}}\,\middle|\,b_{ik}>0,i=1,2,\cdots,m\right\}$$

确定 x_{j_r} 为离基变量．

Step 5　以 b_{rk} 为主元作旋转变换，得新的可行基 $\widetilde{\boldsymbol{B}}$，以 $\widetilde{\boldsymbol{B}}$ 代替 $\boldsymbol{B}$，转 Step 2.

Gilmore 和 Gomory 方法可在如表 12.2.1 所示的单纯形表上进行迭代．

表 12.2.1　Gilmore 和 Gomory 方法迭代表格样式

	x_{j_1}	$\cdots$	x_{j_m}	$x_j(j\in R)$	
x_{j_1}	1	$\cdots$	0	b_{1j}	b_{10}
$\vdots$	$\vdots$		$\vdots$	$\vdots$	$\vdots$
x_{j_m}	0	$\cdots$	1	b_{mj}	b_{m0}
r	0	$\cdots$	0	r_j	

例 12.2.1　用 Gilmore 和 Gomory 方法求解问题（12.1.11）．

解　引入松弛变量 x_3,x_4 和 x_5，我们得到初始基本可行解 $\boldsymbol{x}_0=(0,0,4,6,14)^{\mathrm{T}}$．初始单纯形表如表 12.2.2 所示．

表 12.2.2　Gilmore 和 Gomory 方法初始单纯形表

	x_1	x_2	x_3	x_4	x_5	
x_3	-1	1	1	0	0	4
x_4	0	1	0	1	0	6
x_5	2^*	1	0	0	1	14
r	$-\frac{5}{8}$	$-\frac{1}{8}$	0	0	0	

其中 $\boldsymbol{q}^{\mathrm{T}}\boldsymbol{x}_0+\beta=4$，$\boldsymbol{p}^{\mathrm{T}}\boldsymbol{x}_0+\alpha=2$，$\nabla f(\boldsymbol{x}_0)=\left(-\frac{5}{8},-\frac{1}{8},0,0,0\right)^{\mathrm{T}}$．由于 x_1,x_2

列给出 $B^{-1}N$,故有

$$r(x_N) = (r_1,r_2)^{\mathrm{T}} = \nabla_{x_N} f(x_0) - (B^{-1}N)^{\mathrm{T}} \nabla_{x_B} f(x_0)$$

$$= \begin{bmatrix} -\frac{5}{8} \\ -\frac{1}{8} \end{bmatrix} - \begin{bmatrix} -1 & 0 & 2 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} -\frac{5}{8} \\ -\frac{1}{8} \end{bmatrix}.$$

因 $\max\{-r_1, -r_2\} = \frac{5}{8} = -r_1$,以 x_1 为进基变量,由

$$\theta = \min\left\{\frac{14}{2}\right\} = 7 = b_{51}$$

确定 x_5 为离基变量. 这次迭代结果如表 12.2.3 所示.

表 12.2.3 Gilmore 和 Gomory 方法最优单纯形表

	x_1	x_2	x_3	x_4	x_5	
x_3	0	$\frac{3}{2}$	1	0	$\frac{1}{2}$	11
x_4	0	1	0	1	0	6
x_1	1	$\frac{1}{2}$	0	0	$\frac{1}{2}$	7
r	0	$\frac{62}{121}$	0	0	$\frac{5}{121}$	

当 x_1 代替基中 x_5 时,我们得到点 $x_1 = (7,0,11,6,0)^{\mathrm{T}}$,现令 $q^{\mathrm{T}}x_1 + \beta = 11$, $p^{\mathrm{T}}x_1 + \alpha = -12$,

$$\nabla f(x_1) = \left(-\frac{10}{121}, \frac{47}{121}, 0, 0, 0\right)^{\mathrm{T}},$$

$$r(x_N) = \begin{bmatrix} r_2 \\ r_5 \end{bmatrix} = \begin{bmatrix} \frac{47}{121} \\ 0 \end{bmatrix} - \begin{bmatrix} \frac{3}{2} & 1 & \frac{1}{2} \\ \frac{1}{2} & 0 & \frac{1}{2} \end{bmatrix} \begin{bmatrix} -\frac{10}{121} \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} \frac{62}{121} \\ \frac{5}{121} \end{bmatrix}$$

在表 12.2.3 中,因所有 $r_j \geqslant 0, j \in R$,故 $x_1 = (7,0,11,6,0)^{\mathrm{T}}$ 为最优解. □

12.3 Charnes-Cooper 方法

Charnes 和 Cooper(1962)提出了一种利用变量替换将线性分式规划问题化为线性规划问题来求解的方法. 令

$$t = \frac{1}{q^{\mathrm{T}}x + \beta}, y = tx,$$

则问题(12.1.1)化为

$$\begin{cases} \min & \boldsymbol{p}^{\mathrm{T}}\boldsymbol{y} + \alpha t; \\ \text{s. t.} & \boldsymbol{A}\boldsymbol{y} - \boldsymbol{b}t = \boldsymbol{0}, \\ & \boldsymbol{q}^{\mathrm{T}}\boldsymbol{y} + \beta t = 1, \\ & \boldsymbol{y} \geqslant \boldsymbol{0}, \\ & t \geqslant 0. \end{cases} \tag{12.3.1}$$

定理 12.3.1 如果$(\bar{\boldsymbol{y}},\bar{t})$是问题(12.3.1)的最优解,那么 $\bar{\boldsymbol{x}} = \dfrac{\bar{\boldsymbol{y}}}{\bar{t}}$是问题(12.1.1)的最优解.

证明 首先说明,如果$(\boldsymbol{y},t)$是问题(12.3.1)的任一可行解,那么 $t>0$,若不然,则 $t=0$,从而 $\boldsymbol{y}\neq\boldsymbol{0},\boldsymbol{y}\geqslant\boldsymbol{0},\boldsymbol{A}\boldsymbol{y}=\boldsymbol{0}$. 设 $\boldsymbol{x}$ 是问题(12.3.1)的一个可行解,则对任意常数 $\lambda>0$,有

$$\boldsymbol{A}(\boldsymbol{x}+\lambda\boldsymbol{y}) = \boldsymbol{A}\boldsymbol{x}+\lambda\boldsymbol{A}\boldsymbol{y} = \boldsymbol{b} \text{ 及 } \boldsymbol{x}+\lambda\boldsymbol{y}\geqslant\boldsymbol{0},$$

这说明:$\boldsymbol{x}+\lambda\boldsymbol{y}\in S$,这与 S 为有界集的假设矛盾. 于是,我们有 $\bar{t}>0$.

其次,注意到 $\boldsymbol{A}\bar{\boldsymbol{x}}=\boldsymbol{b},\bar{\boldsymbol{x}}\geqslant\boldsymbol{0}$,因而 $\bar{\boldsymbol{x}}$ 是问题(12.1.1)的可行解.

最后证明 $\bar{\boldsymbol{x}}$ 的最优性. 对任意 $\boldsymbol{x}\in S$,设$(\boldsymbol{y},t)$是问题(12.3.1)与 $\boldsymbol{x}$ 相对应的可行解,则

$$\boldsymbol{y} = \frac{\boldsymbol{x}}{\boldsymbol{q}^{\mathrm{T}}\boldsymbol{x}+\beta}, t = \frac{1}{\boldsymbol{q}^{\mathrm{T}}\boldsymbol{x}+\beta}.$$

因$(\bar{\boldsymbol{y}},\bar{t})$是问题(12.3.1)的最优解,则

$$\boldsymbol{p}^{\mathrm{T}}\bar{\boldsymbol{y}}+\alpha\bar{t} \leqslant \boldsymbol{p}^{\mathrm{T}}\boldsymbol{y}+\alpha t,$$

代入$\bar{\boldsymbol{y}},\boldsymbol{y},t$ 得

$$\bar{t}(\boldsymbol{p}^{\mathrm{T}}\bar{\boldsymbol{x}}+\alpha) \leqslant \frac{\boldsymbol{p}^{\mathrm{T}}\boldsymbol{x}+\alpha}{\boldsymbol{q}^{\mathrm{T}}\boldsymbol{x}+\beta},$$

由 $1=\boldsymbol{q}^{\mathrm{T}}\bar{\boldsymbol{y}}+\beta\bar{t}=\bar{t}(\boldsymbol{q}^{\mathrm{T}}\bar{\boldsymbol{x}}+\beta)$知

$$\frac{\boldsymbol{p}^{\mathrm{T}}\bar{\boldsymbol{x}}+\alpha}{\boldsymbol{q}^{\mathrm{T}}\bar{\boldsymbol{x}}+\beta} \leqslant \frac{\boldsymbol{p}^{\mathrm{T}}\boldsymbol{x}+\alpha}{\boldsymbol{q}^{\mathrm{T}}\boldsymbol{x}+\beta},$$

所以 $\bar{\boldsymbol{x}}$ 是问题(12.1.1)的最优解. □

例 12.3.1 用 Charnes-Cooper 方法求解问题(12.1.11)

解 令 $t=\dfrac{1}{x_1+3x_2+4}$,$\boldsymbol{y}=t\boldsymbol{x}$,则该问题转化为如下线性规划问题

$$\begin{cases} \min & -2y_1+y_2+2t; \\ \text{s. t.} & -y_1+y_2-4t\leqslant 0, \\ & y_2-6t\leqslant 0, \\ & 2y_1+y_2-14t\leqslant 0, \\ & y_1+3y_2+4t=1, \\ & y_1,y_2\geqslant 0, \\ & t\geqslant 0. \end{cases} \tag{12.3.2}$$

用单纯形法求得 $y_1=\dfrac{7}{11},y_2=0,t=\dfrac{1}{11}$是问题(12.3.2)的最优解,于是,问题

(12.1.11)的最优解是 $x_1=\frac{y_1}{t}=7, x_2=\frac{y_2}{t}=0$. □

习题十二

1. 求解下列线性分式规划问题:

$$(1)\begin{cases}\max & \dfrac{1+3x_1+3x_2+2x_3}{1+2x_1+x_2+x_3};\\ \text{s. t.} & 2x_1+5x_2+x_3\leqslant 2,\\ & x_1+2x_2+3x_3\leqslant 3,\\ & x_1,x_2,x_3\geqslant 0.\end{cases}$$

$$(2)\begin{cases}\min & \dfrac{-2x_1+3x_2+5x_3+2}{x_1+2x_2+x_3+2};\\ \text{s. t.} & 2x_1+3x_2+x_3\leqslant 12,\\ & x_1+2x_2\geqslant 12,\\ & x_1+x_3\leqslant 8,\\ & x_1,x_2,x_3\geqslant 0.\end{cases}$$

2. 用原始单纯形法求解线性分式规划时,如何求得它的一个初始基本可行解?

3. 说明原始单纯形法是 Gilmore-Gomory 方法的一种简化形式. 试从既约梯度推出原始单纯形法的最优性充分条件 $\pi_j\geqslant 0, j\in R$.

参考文献

[1] Almogy Y, Levin V. A class of fractional programming problems. Opns. Res. ,1971,9:51 -67.

[2] Avriel M. Nonlinear Programming: Analysis and Methods. Englewood Cliffs, New Jersey:Prentice Hall Inc. , 1976.

[3] Bazaraa M S, Shetty C M. Nonlinear Programming: Theory and Algorithms. New York:John Wiley & Sons, 1979.

[4] Beale E M. Cycling in the dual simplex algorithm. Naval Res. Logist Quart. , 1955,2:269 -276.

[5] Bitran G R, Novaes A G. Linear programming with a fractional functionals. Naval Res. Logist Quart. , 1973,9:22 -29.

[6] Bland R G. New finite pivoting rates for the simplex method. Mathematics of Operations Research, 1977,2:103 -107.

[7] Broyden C G. The convergence of a class of double-rank minimization algorithms: part Ⅰ and Ⅱ. J. Inst. Maths. Applo. , 1970,6:76 -90 and 222 -231.

[8] Charnes A. Optimality and degeneracy in linear programming. Econometrica, 1952,2:160 -170.

[9] Charnes A, Cooper W W. Programming with linear fractional functionals. Naval Res. Logist Quart. , 1962,9:181 -186.

[10] 陈庆华,李建平.线性分式规划原始单纯形算法有限性问题.国防科技大学学报,1993,15(2):66 -71.

[11] Curry H B. The method of steepest descent for nonlinear minimization problems. Quart. Appl. Math. , 1944,2:258 -261.

[12] Dantzig G B. Maximization to linear inequalities. In: Koopmans T C(ed.), Activity Analysis of Production and Allocation. New York:John Wiley and Sons, 1951,339 -347.

[13] Dantzig G B. Linear Programming and Extensions, Princeton, New Jersey: Princeton University Press, 1963.

[14] Dantzig G B, Orden A, Wolf P. Note on linear programming. Pacific J. Math. , 1955,5:183 -195.

[15] Davidon W C. Variable metric method for minimization. Research and Development Report ANL - 5990 (Ref.) U. S. ATOMIC Energy Commission, Argonne National Laboratories, 1959 .

[16] 邓乃扬,等.无约束最优化方法.北京:科学出版社,1982.

[17] D' esopo D A. A convex programming procedure. Naval Res. Logist Quart. , 1959,6:33 – 34.

[18] Fany S C, Puthenpura S. Linear Optimization and Extensions: Theory and Algorithms. Englewood Cliffs, New Jersey: Prentice Hall Inc. , 1993.

[19] Fiacco A V, Mccormick G P. Extensions of SUMT for nonlinear programming: equality constraints and extrapolation. Management Science, 1966,12:816 – 829.

[20] Fiacco A V, Mccormick G P. Nonlinear Programming: Sequential Unconstraints Minimization Techniques. New York: John Wiley & Sons, 1968.

[21] Fletcher R. A new approach to variable metric algorithms. Computer J. , 1970,13:317 – 322.

[22] Fletcher R. A general quadratic programming algorithm. J. Inst. Math. Appl. , 1971,7:76 – 91.

[23] Fletcher R, Powell M J D. A rapidly convergent descent method for minimization. Computer. , J 1963,6:163 – 168.

[24] Fletcher R, Reeves C M. Function minimization by conjugate gradients. Computer Journal, 1964,7:149 – 154.

[25] Goldfarb D. A family of variable metric method derived by variation mean. Math. Comput. ,1970,23:23 – 26.

[26] Hooke R, Jeeves T A. Direct search solution of numerical and statistical problems. J. Assoc. Computer Machinery, 1961,8:212 – 229.

[27] 胡毓达.非线性规划.北京:高等教育出版社,1990.

[28] John F. Extremum problem with inequalities as subsidiary conditions, in: Studies and Essays, Courant Anniversary Volume (Friedricks K D et al ed.). New York: Interscience Publishers, 1948,187 – 204.

[29] Kiefer J. Sequential minimax search for a maximum. Proceeding of the American Mathematical Society, 1953,4: 502 – 506.

[30] Kuhn H W, Tucker A W. Nonlinear programming. In: Proceeding of the Second Berkeley Symposium on Mathematical Statistics and Probability (Neyman J ed.). Berkeley and Los Angeles, California: University of California Press, 1951,481 – 492.

[31] Lemke C E. The dual method for solving the linear programming problem. Naval Res. Logist Quart. ,1954,1:36 – 47.

[32] Levenberger R. A modified Marquardt subroutine for nonlinear least squares AERE. Harwall, Report, No. R – 6799,1971.

[33] 李建平.有界变量线性分式规划的原始单纯形算法.湖南数学年刊, 1995,3:130 – 132.

[34] Luenberger D G. Linear and Nonlinear Programming (2nd Edition). London: Addison – Wesley, 1984.

[35] Martos B. Hyperbolic programming. Naval Res. Logist Quart. , 1964, 11:

135 – 155.

[36] Nelder J A, Mead R. A simplex method for function minimization. Computer J., 1965, 7:308 – 313.

[37] Polyak B T. The method problems. USSR Computational Mathematics and Mathematical Physics, 1969, 9:94 – 112.

[38] Powell M J D. An efficient method for finding the minimum of a function of several variables without calculating derivatives. Computer J., 1964, 7: 155 – 162.

[39] Powell M J D. Nonlinear Optimization. London: Academic Press, 1982.

[40] Rockafellar R T. Convex Analysis. Princeton: Princeton Press, 1970.

[41] Rosenbrock H H. An automatic method for finding the greatest or least value of function. Computer J., 1960, 3:175 – 184.

[42] Shanno D F. Condition of quasi-Newton methods for function minimization. Math. Comput., 1970, 24: 647 – 656.

[43] Spendley W, Hext G R, Himsworth F R. Sequential application of simplex designs in optimization and evolutionary operation. Technometrics, 1962, 4: 441 – 461.

[44] Swarup K. Linear fractional programming. Opns. Res., 1965, 13: 1029 – 1036.

[45] Topkis D M, Veinott A F. On the convergence of some feasible direction algorithms for nonlinear programming. SIAM J. Controls, 1967, 5: 268 – 279.

[46] Wilde D J. Optimum Seeking Methods. Englewood Cliffs: Prentice Hall, 1964.

[47] Wolfe P. On the convergence of gradient methods under constraint. IBM J. Research and Development, 1972, 16:407 – 411.

[48] 应玫茜，魏权龄. 非线性规划及其理论. 北京：中国人民大学出版社，1994.

[49] 俞文魮. 单纯形调优法的收敛性质. 中国科学，数学专刊(I)，1979.

[50] 袁亚湘，孙文瑜. 最优化理论与方法. 北京：科学出版社，1997.

[51] 张干宗. 线性规划. 武汉：武汉大学出版社，1990.

[52] 张建中，许绍吉. 线性规划. 北京：科学出版社，1990.

[53] Zoutendijk G. Method of feasible direction. Amsterdam: Elsevier, 1960.

中英文名词索引

A

B

C

D

E

F

G

H

J

K

L

M

N

P

Q

R

S

T

W

X

Y

Z